



UNIVERSIDADE FEDERAL DA BAHIA
Faculdade de Filosofia e Ciências Humanas (São Lázaro)
Programa de Pós-Graduação de Filosofia

**A MENTE INTERPRETADA:
O Realismo Moderado de Davidson e Dennett**

JULIANA DE ORIONE ARRAES FAGUNDES

SALVADOR - BA
JUL. 2018

UNIVERSIDADE FEDERAL DA BAHIA
Faculdade de Filosofia e Ciências Humanas (São Lázaro)
Programa de Pós-Graduação de Filosofia

**A MENTE INTERPRETADA:
O Realismo Moderado de Davidson e Dennett**

JULIANA DE ORIONE ARRAES FAGUNDES

Tese apresentada como requisito à obtenção do grau de Doutora em Filosofia pelo Programa de Pós-Graduação de Filosofia da Universidade Federal da Bahia.

Orientador: Prof. Dr. João Carlos Salles (UFBA)

Coorientador: Prof. Dr. André Leclerc (UnB)

SALVADOR - BA
JUL. 2018

Ficha catalográfica elaborada pelo Sistema Universitário de Bibliotecas (SIBI/UFBA),
com os dados fornecidos pelo(a) autor(a).

de Orione Arraes Fagundes, Juliana
A Mente Interpretada: O Realismo Moderado de
Davidson e Dennett / Juliana de Orione Arraes
Fagundes. -- Salvador, 2018.
243 f.

Orientador: João Carlos Salles Pires da Silva.
Coorientador: André Leclerc.
Tese (Doutorado - Programa de Pós-Graduação em
Filosofia) -- Universidade Federal da Bahia,
Faculdade de Filosofia e Ciências Humanas, 2018.

1. Filosofia da Mente. 2. Interpretivismo. 3.
Davidson. 4. Dennett. I. Salles Pires da Silva, João
Carlos. II. Leclerc, André. III. Título.



UNIVERSIDADE FEDERAL DA BAHIA
FACULDADE DE FILOSOFIA E CIÊNCIAS HUMANAS
PROGRAMA DE PÓS-GRADUAÇÃO EM FILOSOFIA

Estrada de São Lázaro, 197 – Federação – Salvador – Ba – CEP 40210-730
Tel.: (71) 32836439 //Fax: 32836449 – site: www.ppgf.ufba.br/ e-mail: ppgf@ufba.br

ATA DA DEFESA PÚBLICA DA TESE DE JULIANA DE
ORIONE ARRAES FAGUNDES, REALIZADA NO DIA 01 DE
NOVEMBRO DE 2018.

Ao primeiro dia do mês de novembro, às oito horas, na sala de videoconferência do STI-UFBA, foi instalada a sessão pública para julgamento da Tese Final elaborada pela doutoranda **Juliana de Orione Arraes Fagundes**, intitulada: **“A Mente Interpretada: O Realismo Moderado de Davidson e Dennett”**. Em seguida, o Prof. João Carlos Salles Pires da Silva, deu continuidade aos trabalhos. A Banca julgadora foi constituída pelos seguintes professores: Prof. Dr. João Carlos Salles Pires da Silva (Orientador/UFBA), Prof. Dr. André Leclerc (UNB), Prof^ª. Dr^ª. Sofia Miguens (UNIVERSIDADE DO PORTO), Prof^ª. Dr^ª Karla de Almeida Chediak (UERJ), Prof. Dr. Marco Aurélio Oliveira da Silva (UFBA) e Prof. Dr. Rafael Lopes Azize (UFBA). Passou-se à exposição da doutoranda Juliana de Orione Arraes Fagundes, seguida da arguição dos professores integrantes da banca. A Banca reuniu-se, então, em separado, tendo resolvido, por unanimidade pela sua **Aprovação**. Nada mais havendo a tratar, foi encerrada a sessão e lavrada a presente ata. Salvador, 01 de novembro de 2018.

João Carlos Salles Pires da Silva

Rafael Lopes Azize

João Carlos Salles Pires da Silva

Juliana de Orione Arraes Fagundes

João Carlos Salles Pires da Silva

Dedico aos meus pais.

AGRADECIMENTOS

Há pessoas grandes que doam muito para os pequenos, mesmo sabendo que jamais estaremos à altura de retribuí-los: os professores. A eles devemos nossa inserção no mundo dos humanos e a realização de cada um de nossos sonhos. Eu tenho a sorte de poder incluir nesse grupo de nobres também os meus pais. Além de me darem a vida, criaram condições para que eu pudesse seguir meus passos. Muito obrigada! Tentarei seguir seu exemplo para levar aos meus próprios alunos o conhecimento recebido.

Ao professor André Leclerc, essa figura transbordante de generosidade que me aceitou como orientanda adotiva. A ele deve ser atribuído o mérito por orientar, criticar e sugerir leituras sem interferir no desenvolvimento de meus próprios argumentos, deixando-me livre para traçar um caminho, ainda que houvesse pontos de discordância. Sua orientação me trouxe segurança em meu próprio trabalho, e assim eu pude arriscar. É uma grande honra ter tido essa oportunidade de interação com ele.

Ao professor João Carlos Salles que me aceitou como orientanda, permitindo minha inserção no Programa de Pós-Graduação da Universidade Federal da Bahia. Mesmo tendo muito pouco tempo disponível, por estar dedicado importante função de reitor, sempre se dispôs a me atender quando procurado, oferecendo um apoio indispensável à concretização desta pesquisa.

À professora Maria Eunice Quilici Gonzalez pelo acolhimento. Em um momento em que eu temia não conseguir continuar, uma breve conversa com ela me abriu novos horizontes. A partir daí, pude desatar os nós e buscar soluções para as dificuldades. Sua experiência e paciência significaram muito para mim.

Ao professor Waldomiro Silva Filho pelas sugestões de leituras e envio de material, também por ter contribuído com sua experiência e conhecimento acumulado na fase de elaboração do projeto de doutorado, apontando pontos frágeis.

Ao professor Rafael Lopes Azize pela leitura cuidadosa do texto da minha qualificação e subsequentes questionamentos que me levaram a algumas reformulações. Agradeço também pelas perguntas em minha banca de defesa e pela gentil oferta de fazer uma revisão no “abstract” da tese.

À professora Sofia Miguens pelas breves conversas antes da defesa e pelo material que me enviou, muito útil no desenvolvimento desta tese. Seu trabalho me permitiu alguns “insights” sobre a obra de Dennett. Também agradeço por toda colaboração com a organização da banca de defesa e, em especial, pela leitura cuidadosa e crítica que fez do meu trabalho, trazendo uma série de questões.

À professora Karla Chediak, que também fez uma leitura bem atenta e trouxe uma série de questões, permitindo-me reflexão e esclarecimento acerca dos principais pontos da tese.

Ao professor Marco Aurélio Oliveira da Silva, um agradecimento especial. Além de ler meu trabalho e participar da banca oferecendo dúvidas e fazendo sugestões, ofereceu um importante apoio prático. Foi muito difícil encontrar uma data em que todos os membros da banca estivessem disponíveis, o que chegou a me causar ansiedade. Porém, esse professor me ajudou com cada detalhe prático para essa organização, sua entrada no processo fez que tudo começasse a fluir com leveza.

Ao professor Paulo Abrantes, meu orientador na graduação e no mestrado. grata pelo que me ensinou desde o início de minhas pesquisas em filosofia e pelas discussões sobre esta tese, trazendo-me confiança para seguir. Ele diz que hoje somos colegas e me pede para tratá-lo como igual, mas para mim, ele continua professor.

Ao grupo Mente, Linguagem e Evolução da UnB e ao seu atual coordenador, Cláudio Reis. O grupo me permitiu espaço para apresentar diversas etapas da tese. Também contribuiu com outras discussões e leituras, permitindo um olhar mais amplo sobre meu tema. O grupo foi um companheiro intelectual, ajudou a manter o ânimo ao longo das leituras e da escrita.

Ao GT de Filosofia da Mente da ANPOF, onde também pude apresentar etapas do trabalho e receber críticas e sugestões. É um GT sempre aberto para que os alunos de pós-graduação possam discutir suas pesquisas, um espaço inestimável de troca acadêmica.

À Universidade Estadual do Sudoeste da Bahia, instituição que me permitiu dois anos de licença e uma bolsa para que eu pudesse concluir a tese em outro estado, tendo a oportunidade de interagir com todas essas figuras admiráveis durante o processo.

Sem o apoio de cada pessoa ou instituição acima citadas, este trabalho não existiria. Até aqui, falei os grandes, mas há também os iguais: os meus irmãos, as minhas amigas e o meu namorado. Sem eles, eu não teria a estabilidade emocional para me dedicar às horas de solidão exigidas na escrita de uma tese de doutorado em filosofia. Não posso me esquecer aqui da minha pequena companheira ronronante de todas as horas.

Sou muito grata!

SUMÁRIO

INTRODUÇÃO	13
Principais argumentos interpretivistas de Davidson e Dennett	16
Realismo ou antirrealismo?	19
Perguntas que nortearam a pesquisa	21
1. O PONTO DE PARTIDA	27
1.1. A negação do Teatro Cartesiano	28
<i>1.1.1. A mente descentralizada de Dennett</i>	<i>30</i>
1.2. Pausa para os <i>qualia</i>	34
<i>1.2.1. As cores e as outras mentes</i>	<i>37</i>
<i>1.2.2. Zumbis: uma concepção de qualia a ser revista</i>	<i>38</i>
1.3. A minha mente e as outras mentes	41
<i>1.3.1. Digressão sobre ontologia</i>	<i>45</i>
<i>1.3.2. Fechamento causal do mundo físico e estados mentais causais</i>	<i>48</i>
<i>1.3.3. Caráter anômalo do mental</i>	<i>51</i>
1.4. Um bom reducionismo	54
<i>1.4.1. Sobre o algoritmo evolutivo</i>	<i>58</i>
1.5. Intencionalidade	61
<i>1.5.1 Interpretivismo cotidiano e idealizado</i>	<i>63</i>
1.6 O ancestral do interpretivismo	65
<i>1.6.1. Indeterminação da tradução</i>	<i>70</i>
2. INTERPRETIVISMO DE DONALD DAVIDSON: COMO O	
DISCURSO CONSTITUI UMA REDE DE ESTADOS MENTAIS	75
2.1. Interpretação radical	77
<i>2.1.1. Davidson e a verdade nas linguagens naturais</i>	<i>81</i>
<i>2.1.2. É possível definir a verdade?</i>	<i>84</i>
<i>2.1.3. Tradução e interpretação</i>	<i>87</i>
2.2. Princípio de caridade e o terceiro dogma do empirismo	91
<i>2.2.1. Relativismo ou realismo?</i>	<i>96</i>
2.3. Triangulação	101
<i>2.3.1. Pensando em bases evolutivas para a triangulação</i>	<i>104</i>
<i>2.3.2. Intencionalidade compartilhada e linguagem</i>	<i>111</i>

2.3.3. <i>Razões e verdade</i>	113
2.4 O sujeito intersubjetivo	116
3. INTERPRETIVISMO DE DENNETT: INTENCIONALIDADE	
ATRIBUÍDA	119
3.1. Considerações metodológicas	120
3.1.1. <i>Posturas para explicar e prever o comportamento de um organismo</i>	123
3.1.2. <i>Sistemas intencionais</i>	126
3.1.4 <i>Intencionalidade intrínseca e derivada</i>	128
3.1.5. <i>A postura intencional e o princípio de caridade</i>	131
3.1.6. <i>Intencionalidade da natureza</i>	133
3.1.7. <i>Mentes dentro de mentes</i>	135
3.2. Teoria da mente	137
3.2.1. <i>A Psicologia de Senso Comum</i>	138
3.3. O método heterofenômico	142
3.3.1. <i>O mundo heterofenômico de um robô</i>	146
3.4. Usos diferentes e complementares para o termo	
“consciência”	148
3.4.1. <i>Ilusionismo</i>	150
3.4.2. <i>A mente cultural dos memes</i>	152
3.4.3. <i>O que seria um meme?</i>	155
3.5. Evolução, mente e linguagem em Dennett	158
4. COMO COMPREENDER OS ESTADOS MENTAIS	161
4.1. Davidson: a verdade das atribuições de estados mentais	162
4.1.1. <i>Quais não são as concepções de verdade de Davidson</i>	163
4.1.2. <i>A analogia com os sistemas de medição</i>	165
4.1.3. <i>Indeterminação da interpretação</i>	167
4.1.4. <i>A herança quineana e o caminho davidsoniano</i>	169
4.1.5. <i>A irreduzibilidade tem a ver com a indeterminação?</i>	170
4.1.6. <i>Objetividade da subjetividade</i>	172
4.2. Dennett: a realidade que provém da utilidade	174

4.2.1. Padrões de crenças	177
4.2.2. Dennett critica Davidson acerca da ontologia das atitudes proposicionais	180
4.2.3. Resposta de Davidson a Dennett	182
4.2.4. Moedas como sistemas de medição	185
4.2.5. Dennett e a analogia com o dinheiro	188
4.3. A questão ontológica: Realismo moderado	189
4.3.1. Mais sobre a discussão entre Davidson e Dennett acerca do realismo	191
4.3.2. Uma compreensão mais recente do termo “ontologia” por Dennett	194
4.3.3. Há um tipo de dualismo nesses autores?	197
4.4. A questão epistemológica: Falsas crenças não andam sós	200
4.4.1. Truques cerebrais	206
4.5. Questões de tendência ontológica e questões de tendência epistemológica	207
 CONSIDERAÇÕES FINAIS: O PONTO DE VISTA DE	
TERCEIRA PESSOA	211
Interpretivismo não é eliminativismo	214
Interpretivismo também não é epifenomenalismo	216
Retomada de dois argumentos em favor do realismo moderado	222
Os qualia foram eliminados?	226
Caridade e racionalidade	227
 REFERÊNCIAS	 230

A MENTE INTERPRETADA: O REALISMO MODERADO DE DAVIDSON E DENNETT

Interpretivismo é uma posição em filosofia que busca uma compreensão do mental a partir da perspectiva de terceira pessoa. Dois autores são considerados seus representantes centrais: Daniel Dennett e Donald Davidson. Para que seja possível a comunicação, é necessário haver interpretação comportamental, o que demanda a atribuição de uma série de estados mentais. Em Davidson, há os princípios de caridade e racionalidade. De acordo com ele, é preciso pressupor uma alta taxa de coincidência entre as crenças de alguém e as de seu interlocutor para que eles possam se comunicar. Nessa mesma direção, Dennett argumenta que os comportamentos de determinadas entidades só podem ser explicados e previstos por meio da adoção da postura intencional, isto é, por meio da atribuição de intencionalidade. Essa atribuição nunca é de crenças e desejos isolados, mas de um conjunto difuso de estados mentais, incluindo crenças majoritariamente verdadeiras. Para esses autores, não é possível abrir mão de um vocabulário tipicamente intencional. O vocabulário da psicologia, portanto, não pode ser reduzido ao da física. Por isso, o interpretivismo é considerado um tipo de materialismo não reducionista. Eles defendem que o ponto de vista do intérprete tem prioridade sobre o do sujeito na busca de uma compreensão do mental. Propõem uma inversão no ponto de partida: a pesquisa deve se iniciar de fora para dentro, caso contrário, fica estéril, encarcerada no sujeito. Esta tese tem por objetivo fazer uma apresentação das posições interpretivistas de cada um desses dois autores, investigando se são realistas ou antirrealistas acerca do mental. Nessa investigação, será de extrema relevância compreender o papel epistemológico atribuído pelos defensores dessa posição aos estados mentais. Como resultado temos que, para esses autores, o mental precisa ser compreendido a partir das relações do sujeito com o mundo e com os demais sujeitos. Dennett encontra na história evolutiva as razões que estabelecem essas relações. Davidson as encontra nos processos de triangulação presentes em qualquer situação de comunicação. Argumenta-se, aqui, que a existência dos estados mentais não pode ser negada quando se adota uma perspectiva interpretivista, ainda que esses estados mentais não sejam localizados dentro do sujeito. A razão é que eles possuem um papel epistemológico indispensável. Sem eles, ninguém poderia dar nenhum sentido aos comportamentos alheios, muitos menos se comunicar com os outros. A força ontológica não provém, aqui, de serem evidentes a partir de um ponto de vista de primeira pessoa. A realidade dos estados mentais, numa posição interpretivista, está atrelada à sua utilidade e indispensabilidade. Essa posição tem a vantagem de ser compatível com uma abordagem científica do mental, pois considera-o a partir de um ponto de vista publicamente acessível. Ao mesmo tempo, essa posição reconhece e mantém o papel da psicologia de senso comum, sem ser eliminativista nem reducionista.

Palavras-chave: Filosofia da Mente; Interpretivismo; Davidson; Dennett.

MIND INTERPRETED: THE MODERATE REALISM OF DAVIDSON AND DENNETT

Interpretivism is a philosophical position that seeks to understand the mental from the third person point of view. The main representatives of this position are Daniel Dennett e Donald Davidson. In order for communication to be possible, a behavioral interpretation needs to take place, which demands the attribution of a number of mental states. In Davidson, one can find the principles of charity and rationality. According to him, one must presuppose a high rate of coincidence between someone's beliefs and his interlocutor so that they can communicate. In the same direction, Dennett argues that the behavior of particular actors can only be explained and anticipated through the adoption of the Intentional stance, i.e., through the attribution of intentionality. Such attribution is never of isolated beliefs and desires, but of a diffused set of mental states, including a background of beliefs that are mostly true. To these authors, it is not possible to relinquish of a typically intentional vocabulary. The vocabulary of psychology, therefore, cannot be reduced to that of physics. For this reason, interpretivism is considered a form of non-reductive materialism. They advocate that the standpoint of the interpreter takes precedent over subject's viewpoint in the search for an understanding of the mind. They call for an inversion of the starting-point: research should start from the outside to the inside; otherwise, it becomes barren, imprisoned within the subject. This text aims to present the interpretative headings of those two authors, asking if they are realistic or unrealistic concerning the mind. In this investigation, it is crucial to understand the epistemological role assigned to mental states by supporters of that approach. We claim that, for these authors, the mind have to be understood from the subject's relations to the world and with others subjects. Dennett finds in the evolutionary history the grounds on which those relations are set. Davidson finds them in the triangulation processes present in any situation of communication. We argue that the existence of mental states cannot be denied if you adopt an interpretivist perspective, even if these mind states are not situated within the subject. The reason is that they have a crucial epistemological role. Without them, one could not give any meaning for someone else's behavior, much less communicate with others. The ontological power does not stem, here, from them being obvious from a first-person standpoint. The reality of the states of mind, in an interpretivist position, is bound to its usefulness and indispensability. This position has the advantage of being compatible with a scientific take on the mind, because it approaches the mental from a publicly accessible viewpoint. Simultaneously, such a position acknowledges and sustains the role played by common-sense psychology, being neither eliminativist nor reductionist.

Keywords: Philosophy of Mind; Interpretivism; Davidson; Dennett.

INTRODUÇÃO

A mente humana, gradualmente forjada por um processo longo e poderoso, resulta das interações de nossos ancestrais com o ambiente. Porém, uma vez construída, a mente possibilita diversos modos de pensar sobre a realidade, sendo a filosofia um deles. O pensamento filosófico nasce e se reproduz, formando uma árvore da qual um dos ramos tenta compreender o que somos, o que é a própria mente. Nesse afã, uma estratégia filosófica muito comum foi isolar a mente do restante das coisas, criando a ilusão de que ela está e sempre esteve encarcerada dentro de um crânio¹. Não se deve esquecer, contudo, de sua origem: a mente subjetiva é construída em cada indivíduo por meio de suas interações sociais.

O ser humano precisa ser socializado, precisa aprender com os outros a pensar e a falar. Só depois dessa aprendizagem é que se pode formar a noção de uma individualidade solitária e incompreendida, misteriosa para todos, exceto para si mesma. Talvez haja uma chave para esse mistério se o mental for visto a partir da sua construção em contado com o mundo objetivo e com os demais sujeitos. É possível que, a partir de uma harmonização entre os ambientes externo e interno, se possa esperar um maior esclarecimento do fenômeno mental.

Atualmente, entre as diversas tentativas de explicar o mental de modo harmônico com nossas concepções de natureza, há o interpretivismo, posição atribuída a Donald Davidson e a Daniel Dennett. Thornton (2011) apresenta o interpretivismo como um tipo de materialismo não reducionista, pois do ponto de vista ontológico, estamos diante de um tipo de materialismo, os estados mentais não são algo além do mundo físico. Porém, do ponto de vista epistemológico, defende-se que há mais de um recorte pelo qual a realidade possa ser compreendida e, dependendo do aspecto que se pretenda explicar, há uma forma apropriada de se olhar para ela. Nesse sentido, não seria possível defender um reducionismo do ponto de vista epistemológico. A redução é algo que ocorre entre itens linguísticos, não entre categorias ontológicas. Os princípios que regem o vocabulário mentalista são diferentes dos princípios que regem o vocabulário materialista.

Interpretivismo é uma posição que recebe alguma influência do behaviorismo analítico², defende a possibilidade de acesso externo aos estados mentais de outras pessoas, ou

¹ No decorrer do trabalho, se perceberá que isso talvez não seja apenas uma estratégia filosófica. Para Dennett, essa é uma ilusão caracteristicamente humana.

² De acordo com Graham (2015), behaviorismo analítico é a posição segundo a qual os estados mentais são disposições que se tornam manifestas a partir do comportamento da pessoa em uma situação. É uma teoria acerca do significado dos termos ou dos conceitos mentais. Vale notar que o behaviorismo analítico não é o mesmo que o interpretivismo, pois a primeira posição apenas identifica semanticamente estados mentais com disposições comportamentais. O interpretivismo, por outro lado, exige que possamos atribuir estados mentais a

ao menos às suas atitudes proposicionais³, em condições ideais. A atribuição de crenças e desejos, sob uma perspectiva ideal, seria suficiente para a caracterização dos estados internos de um indivíduo. Apesar da idealização filosófica, o interpretivismo é uma posição que reconhece o papel das atribuições de estados mentais em nossas relações quotidianas com os outros. É das nossas situações de comunicação concretas que parte essa posição. O ponto de vista do intérprete se torna fundamental para que se possam atribuir estados mentais a alguém ou a algo e diz-nos, de fato, se esse alguém ou algo possui estados mentais. Tanto Davidson quanto Dennett foram considerados defensores dessa posição. Por diferentes trilhas, ambos defendem uma transparência dos estados mentais sob a perspectiva de uma terceira pessoa que faz atribuição de estados mentais, ao menos em termos teóricos, isto é, sob condições filosoficamente apropriadas. Uma diferença importante entre o behaviorismo e o interpretivismo é que o behaviorismo geralmente se apoia apenas sobre o comportamento para determinar os estados mentais de um sujeito, mas o interpretivismo possui, além do comportamento, os estados internos atribuídos aos sujeitos e uma série de fatores ambientais e não comportamentais que são relevantes para o estudo do mental. O interpretivismo, afinal, é uma posição fundamentada no externismo⁴ acerca do mental.

A diversidade de possibilidades explicativas deve ser preservada para que possamos encontrar respostas relevantes aos nossos diversos problemas. Isso não significa que a realidade seja diversa ou que haja vários níveis ontológicos, apenas que nós, os intérpretes dessa realidade, temos diversas maneiras de nos aproximarmos dela e, para cada ocasião, há uma maneira mais apropriada. Ao tratar de tipos de reducionismos, Dennett diz o seguinte:

De acordo com as interpretações amenas, é possível (e desejável) *unificar* a química com a física, a biologia com a química e, sim, até as ciências sociais com a biologia. Afinal de contas, as sociedades são compostas de seres humanos que, sendo mamíferos, devem obedecer aos princípios biológicos que incluem todos os mamíferos. Estes, por sua vez, são compostos por moléculas, que devem obedecer às leis da química, que por sua vez devem obedecer às regularidades da física subjacente. Nenhum cientista sensato questiona essa leitura amena; todos os tribunais de justiça juntos encontram-se tão unidos pela lei da gravidade quanto

um sujeito para que ele possa ser considerado de fato um sujeito de estados mentais. Algumas versões de behaviorismo analítico, inclusive, rejeitariam o vocabulário mentalista e exigiriam que ele fosse substituído por um vocabulário relativo ao comportamento público. Para o interpretivismo, ao contrário, não é possível abrir mão do vocabulário mental.

³ O termo “atitude proposicional” designa atitudes do sujeito frente a uma proposição. Por exemplo: “Vanessa crê que *p*”, onde “*p*” é uma proposição qualquer, como “a biblioteca da UnB fecha às 18 horas aos sábados”. Outros exemplos de atitudes proposicionais são “pensar”, “desejar”, “esperar”, “recear” etc. Atitudes proposicionais são, portanto, estados mentais cujo conteúdo pode ser expresso por uma proposição particular, como crença, desejo e dúvida. Para muitos, as atitudes proposicionais se contrastam com os *qualia*, aspectos individuais de nossas experiências subjetivas, supostamente sem qualquer manifestação linguística.

⁴ Para o externismo, nossos conteúdos mentais dependem, para sua constituição, de relações do sujeito com o mundo exterior, seja ele natural ou social. Já o internismo defende que os conteúdos dos estados mentais de um sujeito são completamente internos, dependem apenas do sujeito. (SMITH, 2018). Ver: seção 1.3.

qualquer avalanche, porque também são, no final, uma coleção de objetos físicos. Segundo as interpretações grotescas, os reducionistas preferem abandonar os princípios, teorias, vocabulários e leis das ciências de alto nível usando termos inferiores (DENNETT, 1995/1998a, p.85, grifo do autor).

Na concepção de Dennett, embora a realidade seja uma só, sua organização se torna progressivamente mais complexa, de modo que uma explicação para o mental a partir de uma perspectiva da física (como disciplina) não é capaz de capturar aquilo que há de relevante. Não faz nenhum sentido tentar usar um microscópio para ler um romance. As diversas ciências se sobrepõem como diversos olhares para a realidade. Cada uma delas possui recursos explicativos próprios, necessários conforme o nível de explicação exigido. Davidson parece concordar: “Posso imaginar uma ciência que se ocupe de como pensam e atuam as pessoas e que esteja livre de ‘psicologia de senso comum⁵’, mas não posso ver que interesse teria” (1987, p. 54, aspas internas do original, tradução nossa). É preciso usar os recursos das ciências de nível alto para que se possa compreender o mental.

A abordagem de Davidson acerca do mental é chamada de “monismo anômalo”. Monismo porque não há substâncias ontologicamente distintas, mas anômalo porque, se tentarmos explicar a realidade a partir de uma perspectiva física, tudo se submete a determinadas leis gerais, ao passo que as explicações mentais não se submetem a esse tipo de lei. Nesse sentido, o monismo anômalo de Davidson (1970/1994b) segue pela mesma direção quando afirma seu terceiro princípio. O autor procura conciliar três princípios:

(i) que ao menos alguns eventos mentais possuem causas físicas – *Princípio da interação causal*; (ii) que eventos relacionados como causa e efeito recaem sob leis estritas (isto é, leis que são precisas, explícitas e tão sem exceção quanto possível [“as exceptionless as possible”]) – *Princípio do caráter nomológico da causalidade*; e (iii) que não há leis estritas (em oposição a meras generalizações) relacionando eventos mentais e físicos – *Anomalismo do mental* (MALPAS, 2009; grifos do original, tradução nossa).

Sua posição é monista porque, para ele, não há uma distinção ontológica entre o físico e o mental. Os mesmos eventos podem tanto ser descritos por uma perspectiva dos aspectos mentais quanto dos aspectos físicos, mas a sua natureza é uma só. Cada vez que se descreve um evento mental, é possível fazer uma descrição física desse mesmo evento (o mesmo não vale para todos os eventos físicos), mas tal descrição será de muito pouca utilidade se pretendemos, por exemplo, explicar ou prever o comportamento de um agente.

⁵ Psicologia de senso comum são as teorias quotidianas que as pessoas têm do comportamento umas das outras. Para podermos explicar e prever o comportamento alheio, atribuímos aos outros determinados estados mentais. Fazemos isso quotidiana e naturalmente, sem maiores questionamentos ou reflexões. Ver: seção 3.2.

Assim, de acordo com o monismo anômalo, a psicologia não pode ser reduzida à física, embora possua a mesma ontologia. Davidson desenvolve sua argumentação para mostrar que não é necessário acrescentarmos nada à nossa ontologia da física nem é necessário abandonarmos os nossos conceitos da psicologia de senso comum. Na realidade, se o que queremos é compreender o comportamento humano, a psicologia de senso comum se torna indispensável.

De fato, há diversas perspectivas a partir das quais podemos abordar algo que estejamos buscando conhecer, há vários ângulos por meio dos quais podemos observar algo. Certa vez, eu estava em um museu com um amigo e havia uma tela com uma moldura bastante trabalhada que fazia uma sombra curiosamente detalhada na parede. O amigo que estava comigo olhou para o quadro pela lateral e me revelou ironicamente o que ele mais admirava naquela tela: a sombra que ela projetava na parede. Ora, olhar para a sombra de uma tela pela lateral não diz nada sobre a obra. Para que seja possível apreciar a pintura, é preciso que olhemos para a tela pela frente. Mas o meu amigo havia lançado um olhar diferente à tela a partir de um ângulo improvável e tive a oportunidade de compartilhar com ele a experiência incrível da estranha sombra projetada na parede, dando um sentido novo à apreciação do quadro.

Os diferentes modos de olhar por diversos ângulos e enfoques correspondem às diversas formas de compreender a realidade. É possível dar uma explicação para os comportamentos de um agente fundamentada nas leis estritas das ciências naturais. Contudo, essa explicação costuma ser inútil, poderia até explicar o movimento, mas não explicaria o gesto, pois o gesto tem um significado. Ao tentar compreender o comportamento de alguém, nós geralmente buscamos as razões pelas quais a pessoa está agindo dessa forma. É preciso relacionar aquele comportamento aos estados mentais do sujeito, e apenas assim conseguiremos dar um sentido ao que a pessoa está fazendo. Essa é uma compreensão fundamental para todas as nossas relações sociais, em especial, para a comunicação.

Principais argumentos interpretivistas de Davidson e Dennett

Podemos encontrar em cada um desses autores argumentos que favorecem a perspectiva interpretivista do mental. Os principais, para Davidson, são o princípio de caridade e o desenvolvimento que oferece da noção de intérprete radical. São noções bastante entrelaçadas entre si. Em Dennett, encontramos o método heterofenômico para a

compreensão dos estados internos de um sujeito e a postura intencional, usada para explicar e prever o comportamento de um agente.

Resumidamente, o princípio de caridade consiste na ideia de que, para que se compreenda o comportamento de um agente, precisamos atribuir às crenças dele alto grau de racionalidade, coerência e verdade. Certamente, isso tudo precisa ser atribuído também a nós mesmos, de modo que haverá, ainda, a suposição de um alto grau de coincidência entre nossas crenças e as de quem pretendemos compreender.

Já o intérprete radical é uma espécie de alegoria que esclarece simultaneamente a filosofia da linguagem e a filosofia da mente de Davidson. É alguém que precisa interpretar o discurso de um sujeito estando muito pouco informado acerca do que se passa naquela mente. Porém, para a interpretação poder funcionar, o intérprete terá que fazer uma atribuição mínima de estados mentais ao enunciador (aqui, entra o princípio de caridade) descobrindo, por exemplo, quais são as reações dele diante do que considera verdadeiro ou falso.

Inicialmente, isso envolve um processo de triangulação entre falante, intérprete e mundo, pois é o contato desses dois agentes entre si e com a realidade externa que vai dar sustentação a essa atribuição de estados mentais. Nesse sentido, o argumento da triangulação também ocupa um papel bastante central no interpretivismo do autor. Ao fazer essa atribuição, se o intérprete obtém sucesso na continuidade da comunicação, significa que a atribuição de estados mentais também obteve alta taxa de sucesso. O processo de triangulação se inicia por ostensão, formando um círculo concreto entre falante, intérprete e mundo. Isso forma uma camada básica da linguagem sobre a qual se erguem os aspectos mais abstratos do pensamento e da linguagem, cuja referência não pode ser situada no mundo físico, mas apenas pela internalização de normas sociais mais complexas. Por isso, se não houver a triangulação, não é possível a formação de conceitos.

Pegemos o exemplo de duas pessoas que discutem sobre milagres, sendo que uma delas acredita que existem, e a outra não acredita. Ambas são capazes de compreender uma à outra e apresentar argumentos, ou seja, o fato de discordarem acerca desse ponto não impede o diálogo. Porém, o debate entre elas só é possível porque cada uma acredita que milagres são supostas situações capazes de quebrar com suas expectativas comuns, parecendo desafiar a ordem natural das coisas. A diferença é que uma delas acredita que isso é mais do que uma aparência. Ela poderia defender, por exemplo, que há forças sobrenaturais interferindo em nossas vidas, mas ambas precisariam concordar que o “sobrenatural” não pode ser conhecido pelas mesmas vias experimentais e racionais com que alcançamos nossas informações cotidianas. Precisarão também compartilhar muitos outros pontos de acordo (ainda que haja

divergências) sobre como tende a ser a ordem natural das coisas, sobre como as conhecemos, sobre o que significam as palavras que usamos, sobre o fato de que estão em diálogo, sobre estarem em desacordo acerca dos milagres, sobre o fato de que o interlocutor está tentando dizer algo com sentido, sobre estarem falando línguas mutuamente compreensíveis, ou a mesma língua. Enfim, será necessário elas compartilharem um fundo massivo de experiências compartilhadas. Porém, não é sobre as crenças compartilhadas que eles conversam. É desnecessário conversar sobre elas. Elas são o fundamento tácito necessário à compreensão do que o outro diz.

A razão pela qual nós não podemos entender o que um homem quer dizer sem saber bastante sobre suas crenças é essa. Para interpretar o comportamento verbal, devemos ser capazes de dizer quando um falante sustenta uma frase como verdadeira. Mas as frases são consideradas verdadeiras parcialmente por causa do que se crê e parcialmente por causa do significado que as palavras do falante têm (DAVIDSON, 2001, p. 338, tradução nossa).

Por isso, todo comportamento de um agente só pode ser compreendido se lhe atribuirmos uma série de estados mentais. E aqui, há uma via de mão dupla. Ao obtermos sucesso na compreensão desse comportamento, por exemplo, se conseguimos dar sentido ao comportamento alheio e nos comportamos de modo a sustentar a interação, isso significa que muitos dos estados mentais se revelaram apropriadamente para nós. Não é que a mente alheia se torne um livro aberto diante dos nossos olhos, mas também não é um diário de adolescente trancado com cadeado. Algumas páginas se abrem e, lidas com cuidado por um intérprete atento, fazem sentido. Por outro lado, não fariam qualquer sentido se não houvesse um leitor alfabetizado naquela língua. Assim como um livro não faria nenhum sentido se fosse escrito em uma língua conhecida apenas por uma pessoa, também não faz sentido uma mente encarcerada no indivíduo.

Passemos os olhos sobre os argumentos de Dennett. Falando de forma bem geral, adotar a postura intencional é atribuir atitudes proposicionais a um sistema qualquer com o intuito de prever e explicar seu comportamento. Por exemplo, ao jogar xadrez com um computador, é preciso lhe atribuir o desejo de ganhar e a inteligência para realizar os melhores movimentos possíveis com essa meta. Tentar compreender o que ocorre na máquina a partir dos seus circuitos eletrônicos será infrutífero, despenderá muito tempo e não permitirá prever o seu próximo movimento, pois essa compreensão envolve convenções sociais e culturais muito além dos eventos eletrônicos. É possível adotar diversas posturas diante de um sistema a cujo comportamento se pretende dar um sentido, mas se esse sistema for intencional,

uma compreensão eficaz passará pela atribuição àquele sistema de intencionalidade⁶. O que caracteriza um sistema intencional é o sucesso decorrente da adoção da postura intencional, ou seja, a adoção de comportamentos apropriados que permitem a continuidade da interação com aquele sistema.

Já o método heterofenômico é uma proposta para o estudo da vida subjetiva a partir de uma perspectiva de terceira pessoa. Consiste em ouvir os relatos do sujeito acerca de sua própria mente, interpretando-os, com vistas à construção de uma compreensão daquele sujeito. Como essa construção é feita pelo intérprete, ela é denominada heterofenômica. Por esse método, o pesquisador permite que o sujeito constitua um universo subjetivo como uma obra de ficção, mantendo-se agnóstico em relação a se há ou não referência no mundo interno daquele sujeito para as crenças manifestas. Esse agnosticismo, à primeira vista, parece levar Dennett a um grau de antirrealismo acerca dos estados mentais. Em alguns momentos o autor parece não estar interessado na questão acerca de se os estados mentais existem ou não, preocupando-se mais com a importância da atribuição de estados mentais para podermos prever e explicar o comportamento de um sistema. Por essa razão, surge a pergunta: *o interpretivismo é antirrealista acerca dos estados mentais?*

Realismo ou antirrealismo?

John Heil (1998) compreende esse tipo de abordagem como antirrealista. As mentes, nesse sentido, seriam espécies de construtos usados por nós para atribuir estados mentais aos outros e a nós mesmos, mas não aspectos ontológicos da realidade. Segundo ele, a compreensão que Davidson e Dennett têm do mental pode ser comparada a um sistema de coordenadas usado para que possamos conhecer um terreno, mas que não faz parte da própria realidade. “Atribuir pensamentos a agentes, nessa concepção, seria como atribuir uma latitude e uma longitude a um lugar na superfície da Terra” (HEIL, 1998, p. 167). Segundo esse autor, numa abordagem interpretivista, as mentes seriam apenas artifícios usados por nós para que possamos compreender o comportamento dos outros agentes, sem nenhum tipo de referência na realidade. Podemos, contudo, nos perguntar se a abordagem interpretivista é de fato antirrealista.

Vejam, primeiramente, o caso de Dennett. O autor faz questão de diferenciar o interpretivismo, posição que aceita, do interpretacionismo⁷, posição que rejeita. Por um lado,

⁶ A noção de intencionalidade é apresentada na seção 1.5, e o tema será melhor desenvolvido no capítulo 3.

⁷ Em inglês: “interpretivism” e “interpretationism”.

Dennett (1987, p. 14) recusa a ideia considerada por ele como “realismo radical” de que seja possível confirmar a atribuição de crenças encontrando algo dentro da mente do crente. Por outro lado, ele recusa também a ideia de que a interpretação das crenças de alguém exigiria uma análise hermenêutica completamente externa ao crente, sendo essa a posição chamada por ele de “interpretacionismo”. Segundo esta posição, portanto, as crenças seriam apenas atribuições culturais. De acordo com Dennett, o realismo radical daria aos estados mentais um status parecido com o de um vírus, isto é, algo que tem uma existência perfeitamente concreta dentro do sujeito. O interpretacionismo, ao contrário, daria às crenças um status parecido ao de ter estilo, ou seja, algo que só pode ser atribuído a alguém sob uma perspectiva cultural e não tem qualquer existência independente disso. Dennett (1987, p. 15) se propõe a defender uma posição intermediária ao realismo e ao interpretacionismo, na qual a crença é distinguida a partir do ponto de vista de quem adota uma estratégia perante outrem com o intuito de explicar e prever seu comportamento:

Minha tese será que, ao mesmo tempo em que a crença é um fenômeno perfeitamente objetivo (o que aparentemente faz de mim um realista), ela só pode ser discernida a partir do ponto de vista de quem adota certa estratégia preditiva, e sua existência pode ser confirmada apenas por uma avaliação do sucesso dessa estratégia (o que aparentemente faz de mim um interpretacionista) (DENNETT, 1987, p.15, tradução nossa).

A preocupação de Dennett é de buscar uma abordagem do mental que seja compatível com os conhecimentos desenvolvidos pelas diversas ciências. Nesse sentido, para ele, a única possibilidade é explicá-lo a partir do ponto de vista externo, ou seja, o ponto de vista do intérprete. Isso não significa uma negação da existência dos estados mentais, mas sim um deslocamento da posição a partir da qual se pode chegar a eles. Provavelmente, é por isso que Dennett atribui objetividade às crenças, por elas se manifestarem a partir de um ponto de vista externo, que é o ponto de vista apropriado para a apreciação daquilo que é objetivo.

No caso de Davidson, a situação é parecida. Os estados mentais estão lá, mas assim como são construídos a partir do contato direto com o mundo exterior e com os outros seres portadores de mentes, também poderão se revelar a partir da perspectiva do intérprete. Essa posição é denominada por Malpas (2005) “realismo cotidiano”, conforme será desenvolvido⁸. A própria compreensão da noção de objetividade depende tanto da subjetividade quanto da intersubjetividade (DAVIDSON, 1991/2001f).

Essa compreensão talvez ajude a entender como se dá a inserção do mental no mundo material. Dentro da argumentação davidsoniana, o pensamento é linguístico e a linguagem se

⁸ Ver: seção 2.2.1.

constrói no contato direto com a realidade física e social. No mais, é parte do que constitui um sujeito, e os sujeitos estão no mundo, junto aos demais sujeitos e às outras coisas. O fato de que seus estados mentais se tornam transparentes para o intérprete não faz com que o sujeito deixe de ser uma autoridade acerca de seus estados mentais, nem leva Davidson ao antirrealismo.

Byrne (1998) apresenta o interpretivismo como a ideia de que os conteúdos mentais podem, *a priori*, ser capturados por um intérprete ideal. Dizer que um sujeito tem uma crença é dizer que há um intérprete ideal disposto a atribuir a esse sujeito a crença. Segundo esse autor, tal intérprete seria capaz de descobrir crenças e desejos de um sujeito da maneira como elas de fato são. O interpretivista considera que esses estados mentais são completamente acessíveis a partir de um ponto de vista externo por um intérprete ideal e não têm uma existência localizada dentro da mente do sujeito. Em outras palavras, os estados mentais possuem uma existência objetiva, mas não é necessário “abrir a cabeça” do sujeito para que se possam localizá-los. De acordo com Byrne, a formulação da ideia do intérprete ideal é pressuposto para que o interpretivismo seja desenvolvido. Tal intérprete partirá de evidências físicas e comportamentais para que possa fazer suas atribuições de estados mentais.

O interpretivismo consiste em levar a sério as atribuições externas de estados mentais. A própria auto atribuição de estados mentais, em um primeiro momento, se dá durante o processo de socialização da criança⁹. Essa criança precisa compreender o comportamento do mundo que a cerca e o faz com a ajuda de outros seres humanos, atribuindo estados mentais aos outros e a si mesma simultaneamente. Portanto, a atribuição de estados mentais bem sucedida, no mínimo tem muito a dizer acerca do que se passa na mente das outras pessoas, além de ser uma peça sem a qual não poderíamos entender a linguagem nem poderíamos nos fazer compreender pelos outros.

Perguntas que nortearam a pesquisa

No início da pesquisa, foram estabelecidas algumas perguntas que norteariam as investigações. Porém, com o avanço dos estudos, percebeu-se que a abordagem interpretivista se constrói sobre fundamentos que acabam por não abrir espaço para algumas dessas questões¹⁰. Também foram elaboradas hipóteses, sendo que algumas foram corroboradas,

⁹ Aqui, tratamos da história de vida de um indivíduo. É possível supor, contudo, que um processo análogo tenha ocorrido em nossa história evolutiva.

¹⁰ Ver: seção 4.5.

outras não. O fato, contudo, é que ter essas questões e hipóteses como norte foi fundamental para que se pudesse compreender o interpretivismo de Davidson e de Dennett, seus fundamentos e o que essa posição propõe de diferente para um estudo do mental. Por isso, vejamos aqui que perguntas foram essas. As respostas serão deixadas para o final da tese.

Considerando que para o interpretivismo, (1) a perspectiva do intérprete é uma postura plenamente legítima para a compreensão dos estados mentais; (2) as atribuições de estados mentais a um agente por parte de um intérprete revelam os estados mentais daquele agente e (3) há várias descrições igualmente apropriadas para a mesma realidade (assim como há várias interpretações corretas para os mesmos comportamentos), isso significa que essa posição compreende os estados mentais apenas como atribuições externas? Eles não seriam reais ou há algum sentido em que possam ser considerados reais?

O interpretivista atribui crenças a um sujeito para explicar e prever seu comportamento. Também precisa atribuir crenças aos outros para construir sua própria relação com o mundo e consigo mesmo. Em outras palavras: o intérprete também possui crenças, e essas crenças são construídas de forma linguística, a partir de sua relação com a realidade, incluindo a realidade social. O discurso do sujeito e o seu comportamento em geral são o que permite a interpretação e a atribuição de crenças a ele. Por isso, aquilo que é proferido por alguém talvez tenha algum lastro, se não em seu ambiente interior, nas relações que esse sujeito possui com o mundo exterior e os outros sujeitos. Da mesma maneira, qualquer comportamento seria capaz de revelar algo que se passa naquele sujeito.

Outro aspecto que parece relevante é a questão de nossa história evolutiva, como ela se conduziu até nós, seres capazes de uma linguagem convencional e de pensamentos proposicionais? O interpretivismo talvez possa ser corroborado quando adotamos uma perspectiva evolucionista, que é uma perspectiva bastante cara a Dennett. Davidson, embora mencione essa perspectiva, não a explora, mas encontramos em um dos recentes trabalhos do psicólogo evolutivo Michael Tomasello (2014) interessantes “insights” capazes de trazer suporte empírico ao trabalho de Davidson.

Diante disso, há dois grupos de perguntas que se pretende colocar na tese de doutorado a ser desenvolvida: perguntas de tendência ontológica e perguntas de tendência epistemológica, sendo que esses grupos se relacionam. As questões do primeiro grupo orbitam em torno dessa: numa perspectiva interpretivista e, para cada um desses autores, os estados mentais têm uma existência no sujeito ou são apenas algo que se atribui de uma perspectiva externa sem que haja nada além disso? Podemos apresentar essa questão de maneira um pouco mais ampla: as atitudes proposicionais são reais? Têm um status parecido

com a de um sistema de medição, como proposto por Heil (por exemplo, podemos medir a temperatura em graus célsius, em graus fahrenheit, ou poderíamos agora inventar um novo sistema de medição), ou são algo mais consistente do que isso?

Essa mesma pergunta pode se estender aos estados mentais não proposicionais. Embora a posição interpretivista esteja focada apenas nos estados mentais proposicionais, este trabalho pretende abordar a questão da história evolutiva como algo que construiu a mente humana. Nesse sentido, deve-se levar em conta que animais pré-linguísticos estabeleceram determinadas relações com seu entorno e essas relações criaram o solo sobre o qual a linguagem se ergueu. Assim, cabem-nos as perguntas: Há estados mentais pré-linguísticos? Uma postura interpretivista seria capaz de reconhecer tais estados?

A pergunta central do segundo grupo de perguntas é: Numa perspectiva interpretivista e, para cada um desses autores, é possível estabelecer uma relação entre os estados mentais e o mundo? Há alguma tendência das nossas crenças a fornecer informações verdadeiras acerca da realidade? Colocando essa questão de maneira mais ampla: as crenças estabelecem alguma relação epistemológica do organismo com o mundo e/ou os outros sujeitos? Certamente, existem crenças falsas e somos constantemente surpreendidos com enganos relativamente ao que acreditamos, isso deve ser acomodado se queremos compreender que tipo de relação elas estabelecem entre sujeito e mundo. Mas, de um modo geral, há algum grau de confiabilidade nas crenças que atribuímos a nós mesmos e aos outros? Se a resposta à primeira pergunta for não para ambos os autores, talvez não faça sentido partir para a segunda pergunta, pois como poderia uma crença sem existência vincular o ser humano à realidade?

Em relação às hipóteses previamente elaboradas para conduzir a investigação, pensou-se que a resposta à primeira pergunta seria *sim* para ambos os autores, embora isso não estivesse claro em relação a Dennett. No caso de Davidson, as atitudes proposicionais podem ser reconstruídas como disposições causais, isso parece suficiente para dizer que sejam reais. Por outro lado, a resposta à segunda pergunta parecia inicialmente ser *sim* apenas para Davidson. Não é que Dennett defenda a falsidade de nossas crenças em geral, mas ele parecia não estar comprometido com estabelecer uma relação entre as crenças e a realidade. Por isso, sua posição seria chamada interpretivismo instrumentalista e a posição de Davidson, instrumentalismo realista.

Como mencionado acima, ao final das investigações do presente trabalho de pesquisa nem todas as hipóteses foram corroboradas. As ideias de Dennett e de Davidson se revelaram mais sutis do que pareciam quando da elaboração das hipóteses. Instrumentalismo e realismo acabaram por se misturar, e a noção de sujeito se expandiu pelo mundo objetivo. Esses

autores provocam inversões em nosso modo de olhar para o mental. A partir deste ponto, lançadas essas perguntas e hipóteses, elas serão deixadas em segundo plano, para que as respostas não se revelem muito precocemente. Serão respondidas com suas sutilezas e detalhes ao longo do desenvolvimento do trabalho, e ao final do capítulo 4 haverá apenas uma retomada bem resumida das questões aqui apresentadas.

Neste trabalho, as percepções que Dennett e Davidson possuem do interpretivismo serão desenvolvidas e defendidas. Para isso, será seguida uma trilha na qual se procuram compreender os aspectos relevantes dos pensamentos desses dois grandes filósofos na constituição da perspectiva interpretivista. A tese se constitui de quatro capítulos. Para cada um dos autores trabalhados, há um capítulo específico, o segundo e o terceiro capítulos. O primeiro capítulo tenta constituir alguns pontos de partida para que a trilha possa ser percorrida. O último capítulo coloca os dois autores em diálogo, tentando estabelecer o peso ontológico que eles dão aos estados mentais e às crenças.

Durante o primeiro capítulo, os fundamentos para o pensamento interpretivista serão abordados. Será apresentada a noção de Teatro Cartesiano, à qual o interpretivismo procura se opor. Em seguida, uma alternativa a essa concepção oferecida por Dennett. Além disso, a questão dos *qualia* – estados mentais não proposicionais – será enfrentada, pois uma crítica comum às abordagens de Dennett e de Davidson é que eles não abordam o que, para muitos, é o aspecto central a ser tratado pela filosofia da mente. Argumentar-se-á que essa noção, embora pareça clara e intuitiva, é confusa e paradoxal e que uma abordagem a partir dos aspectos relacionais da consciência pode ser mais frutífera.

Será tratada brevemente também a questão da causalção mental e os argumentos apresentados por Davidson e por Dennett contra uma abordagem epifenomenalista. O monismo anômalo de Davidson será um pouco desenvolvido, assim como sua abordagem ao problema das outras mentes, na qual busca salvaguardar a autoridade do sujeito acerca dos próprios estados mentais. Também no primeiro capítulo, serão apresentadas algumas ideias importante de Sellars e de Quine para o desenvolvimento do interpretivismo. Será buscada ainda uma medida apropriada para um reducionismo cuidadoso e não eliminativista. Nesse contexto, a proposta do algoritmo evolutivo de Dennett será apresentada.

O segundo capítulo tratará das principais teses de Davidson em torno do interpretivismo, em especial, o princípio de caridade e a triangulação. Pode-se dizer que, na

concepção de Davidson, para que seja possível compreender um discurso, é preciso atribuir ao interlocutor alto grau de racionalidade, incluindo crenças majoritariamente verdadeiras e compartilhadas com o intérprete. A triangulação nos mostra que essas crenças atribuídas de fato possuem conexão com o mundo compartilhado entre intérprete e falante. Usaremos um dos trabalhos recentes de Tomasello (2014) para conferir um apoio empírico ao trabalho de Davidson. Tomasello o cita algumas vezes em seu livro, trazendo-nos a percepção de que filosofia e ciências podem dialogar e se apoiar mutuamente.

No terceiro capítulo, serão desenvolvidas duas ideias de Dennett relevantes para a sua versão do interpretivismo: a postura intencional e o método heterofenômico. Falando de forma bem geral, adotar a postura intencional é atribuir atitudes proposicionais a um sistema qualquer com o intuito de prever e explicar seu comportamento, como faz a pessoa que joga xadrez contra uma máquina. Já o método heterofenômico é uma proposta para o estudo da vida subjetiva a partir de uma perspectiva de terceira pessoa. Consiste em ouvir os relatos do sujeito acerca de sua subjetividade, interpretando-os, com vistas à construção de uma compreensão daquele sujeito. Como essa construção é feita pelo intérprete, ela é denominada heterofenômica. Outro tema enfrentado no terceiro capítulo é a concepção de consciência de Dennett. Sua postura interpretivista possui consequências importantes para se pensar essa questão, tema central na obra do autor.

O quarto capítulo visa pesquisar a concepção de mente proporcionada pelas perspectivas interpretivistas Davidson e de Dennett. Defender-se-á que esses autores são moderadamente realistas acerca dos estados mentais. Para isso, será necessário enfrentar a questão da indeterminação da interpretação, algo que pode parecer um desafio para uma abordagem realista. A questão das falsas crenças também precisará ser encarada. Este capítulo tem ainda os problemas de ordem ontológica e epistemológica acima apresentados para lidar, a saber: (1) Numa perspectiva interpretivista as crenças são reais? (2) Nessa mesma perspectiva, elas tendem a produzir uma relação confiável entre sujeito e mundo? Na abordagem interpretivista, os estados mentais deixam de ser os constituintes internos do sujeito e passam a ser identificados externamente, a partir da perspectiva do intérprete de um discurso. Além disso, na perspectiva interpretivista, há uma razão para que se atribuam estados mentais a alguém: a razão é a necessidade de explicar e prever o comportamento (inclusive o comportamento linguístico) daquele ser. Nesse sentido, então, os estados mentais seriam apenas instrumentais ou seria possível atribuir a eles uma existência real? A perspectiva interpretivista, contudo, não permite que olhemos para cada um desses problemas isoladamente, eles estão imbricados e são inseparáveis.

Recentemente, Dennett (2017b) publicou um livro no qual, como em outros livros, traz uma abordagem da consciência a partir da história evolutiva. “O livro tem uma estrutura histórica, levando-nos do mundo pré-biótico às mentes humanas e à civilização humana” (NAGEL, 2017, p. 1, tradução nossa). Como de costume, incorpora em suas pesquisas o que de mais recente tem sido desenvolvido nas ciências cognitivas. Nesse livro, ele trata da noção de “ontologia” de um jeito muito especial. Ele parte da distinção de Sellars entre imagem manifesta e imagem de ciência e amplia a noção de imagem de manifesta aos outros organismos vivos. Todos os organismos possuem sensores biológicos para capturarem determinados aspectos do ambiente. Esses aspectos são as “affordances” daquele organismo, isto é, partes do ambiente que se tornam manifestas para cada organismo. A ontologia, segundo essa concepção, é constituída pelos aspectos do mundo que cada organismo alcança. Uma vez que estamos lidando com questões de ordem ontológica e epistemológica, será importante adentrar um pouco nos argumentos desse último livro.

1. O PONTO DE PARTIDA

Donald Davidson e Daniel Dennett são autores pertencentes a uma corrente que busca superar as contradições do dualismo. Recebem influência do behaviorismo, mas estão atentos às suas dificuldades. Além disso, foram ambos alunos de Quine, permanecem fiéis à fonte em muitos pontos. São ramos de um tronco em cuja base estão Skinner, Wittgenstein e Sellars. Os argumentos desses autores vão contra as concepções tradicionais de subjetividade e tendem a favorecer uma leitura dos estados mentais com base na sociabilidade. Dennett e Davidson entram nesse fluxo para desenvolver a concepção interpretivista.

Este capítulo se propõe a estabelecer um ponto de partida, esclarecendo o tipo de concepção que se pretende combater e algumas noções básicas que servem como norte. A ideia é apresentar alguns argumentos capazes de preparar o terreno para que as ideias propostas por esses autores não pareçam estranhas. Ao deixar de lado uma tradição subjetivista e tentar seguir por outra trilha, esses autores acabam por evitar algumas complicações céticas. Isso pode fazer com que suas posições pareçam ingênuas demais perante uma tradição de constante confronto com os argumentos céticos. Essa é mais uma razão pela qual o estabelecimento do ponto de partida é importante. É preciso compreender que esses autores estão rejeitando como fundamento algo ainda muito impregnado na contemporaneidade: o subjetivismo.

O capítulo está estruturado em seis seções. A primeira busca estabelecer de modo geral que tipo de posição está sendo combatida por esses autores e os modelos alternativos a essa posição propostos por Dennett. A segunda seção busca enfrentar precocemente a crítica de que o interpretivismo não trata do que há de mais importante acerca do mental. Pretende-se mostrar que a crítica não é avassaladora como pareceria. Na terceira se falará de como Davidson busca conciliar a sua posição com a autoridade de primeira pessoa, enfrentando, assim, o ceticismo acerca das outras mentes e uma possível lacuna explicativa entre estados mentais intrínsecos e relacionais. Na quarta seção se apresenta uma introdução ao conceito de intencionalidade, pois ele terá um papel importante ao longo do trabalho, principalmente na compreensão do interpretivismo de Dennett. Na quinta seção, haverá algumas palavras sobre o mito de Jones apresentado por Sellars na década de 1950, onde há um mote para o interpretivismo se desenvolver. Também se tratará da indeterminação da tradução, um conceito proposto por Quine que é muito importante para o interpretivismo. A sexta seção fala

de tipos de reducionismo, argumentando que é interessante compreender o mental como algo que se pode explicar por meio de um vocabulário das ciências de nível mais alto. Tentar usar um vocabulário da física para compreendê-lo conduzirá o pensamento apenas a frustrações, pois é um vocabulário restrito demais para um fenômeno tão complexo. No entanto, isso não implica nem que o mental seja algo além da física nem que seja inexplicável sob qualquer perspectiva objetiva.

1.1. A negação do Teatro Cartesiano

A ideia de que as experiências mentais são algo que se apresenta perante o sujeito ou está dentro dele se tornou bastante arraigada dentro do pensamento filosófico moderno, deixando marcas na história da filosofia. Essas marcas ainda se fazem presentes no pensamento contemporâneo, tanto entre filósofos quanto entre cientistas cognitivos, muitas vezes orientando as interpretações dos resultados dos experimentos em psicologia científica. Porém, desde meados do século XX, alguns autores têm se colocado contrariamente a esse tipo de perspectiva, tendo em vista os problemas a que ela conduz. Nessa linha, encontramos Davidson e Dennett.

Desde o início de sua carreira, Dennett (1969) está empenhado em estudar a questão da consciência por uma via que evite o dualismo. É fato que encontramos na filosofia da mente contemporânea autores - por exemplo: Chalmers (1996), McGinn (1989), Kirk (1974) - cuja herança cartesiana é nítida e assumida, embora em geral não sejam dualistas de substância. Porém, segundo Dennett (1991b), há muitos que, mesmo rejeitando o dualismo, ainda estão seduzidos pela ideia de um lugar onde as experiências conscientes se apresentam para o sujeito em um momento determinado. Esse modelo do mental é denominado por ele de “Teatro Cartesiano”, uma espécie de palco onde o mundo interior seria exibido para o expectador: o sujeito consciente. O título “Teatro Cartesiano” está mais ligado, portanto, à herança moderna do que ao pensamento do próprio Descartes. O Teatro Cartesiano defende que há experiências anteriores à linguagem às quais o sujeito possui acesso privado e imediato. Frequentemente também considera que essas experiências são a fonte da subjetividade e que estão na base de todo o nosso conhecimento.

Para Blackmore (2002), alguns autores que tentaram abandonar a ideia do Teatro Cartesiano não conseguiram deixar de lado a noção de fluxo da consciência. As experiências desse fluxo estariam disponíveis para o sujeito, contrariamente a conteúdos inconscientes, os quais não conseguiriam fazer parte de tal fluxo. Mas essa ideia também exigiria uma espécie

de observador central no cérebro para quem as experiências fariam seu show, ou seja, tais autores não conseguiram abandonar o Teatro Cartesiano¹¹. Esse modelo está tão saturado na filosofia contemporânea que é realmente difícil sair dele.

Tradicionalmente, considera-se a introspecção como o único caminho para o sujeito poder ter acesso à sua vida mental. Porém, isso isola o sujeito da realidade externa, tornando seus estados mentais inacessíveis a qualquer compreensão objetiva e dando origem a ceticismos, tanto acerca do mundo exterior quanto acerca das outras mentes. Esse tipo de subjetivismo, na concepção de autores como Dennett e Davidson, são um desvio para o pensamento filosófico, conduzindo-o a equívocos¹². Isso poderia ser evitado por meio de um exercício paciente e persistente de revisão e mudança dos argumentos filosóficos fundados na noção de Teatro Cartesiano.

A nossa própria estrutura cognitiva, segundo Dennett (1991b), talvez favoreça a concepção do mental como um teatro. Temos quotidianamente a necessidade de atribuir estados mentais às outras pessoas para podermos dar sentido ao seu comportamento. Nessa tarefa, costumamos singularizar a mente do outro. É provável que essa singularização facilite o nosso trabalho de intérpretes do comportamento alheio. Mas, uma vez que a interpretação é possível, não faz sentido dar um salto e concluir que a mente do outro e a nossa mente são tão singulares a ponto de estarem fechadas a qualquer interpretação externa. Ainda que as outras pessoas tenham pensamentos que nós não possamos conhecer, isso não conduz necessariamente ao fechamento do conjunto total de pensamentos de uma pessoa dentro dela mesma.

Uma versão bastante comum do Teatro Cartesiano na mentalidade contemporânea seria a que compreende a mente como cérebro. A ideia seria que o cérebro reuniria toda a vida mental de uma pessoa, constituindo-a. Segundo Dennett (2013, pp. 86-87), essa ideia pode ser proveniente da intuição de que, tirando certas partes do corpo, a pessoa continua sendo quem é, mas retirando o cérebro, não sobra a pessoa. Dennett, ao contrário, entende a mente como algo descentralizado. Ele vê na internet uma metáfora interessante para ilustrar esse ponto.

¹¹ Essa crítica de Blackmore, na realidade, é direcionada a Dennett, um importante influenciador do trabalho dela. Ela tenta radicalizar o pensamento dele ao negar o fluxo de consciência.

¹² Além disso, há um problema lógico: a falácia do homúnculo. Se há um local no cérebro onde as experiências se projetam, então, haverá a necessidade de uma mente para recebê-las, uma espécie de homúnculo no cérebro. Porém, esse homúnculo também precisará ter em seu cérebro um lugar para a projeção das experiências, o que nos levaria para uma regressão *ad infinitum*. Dennett (1991b, p. 162), contudo, destaca que isso não ocorre quando se adota um modelo de funcionalismo homuncular, ou o modelo do pandemônio, isto é, a ideia de que o comportamento dos animais ou outros sistemas intencionais é resultado de uma série de eventos muito simples e estúpidos ocorrendo simultaneamente naquele organismo, como se houvesse uma anarquia de demônios agindo concomitantemente. O resultado da combinação dessa multiplicidade de eventos simultâneos, na visão do autor, é um comportamento inteligente. Ver: seção 1.1.1.

Assim como a internet é descentralizada, para ele, nós também somos. Não somos nossos cérebros. “Seu cérebro não entende inglês, você sim” (DENNETT, 2013, p. 88, tradução nossa). Ora, para entender inglês, dependemos de professores e toda uma estrutura social a qual não está no cérebro.

O Teatro Cartesiano, portanto, assume diversas versões. Pode estar no dualismo de substâncias ou no dualismo de propriedades, pode estar no materialismo forte que tenta reduzir o cérebro à mente e até mesmo em versões de behaviorismo. Sempre que se tenta demarcar o mental sem levar em conta seu aspecto intersubjetivo, abrem-se as cortinas do Teatro Cartesiano. Por exemplo, se uma versão de behaviorismo tenta reduzir o mental apenas aos comportamentos observáveis, estes se tornam os atores em cena perante o behaviorista.

Davidson é um autor que também pretende rejeitar toda forma de subjetivismo, buscando compreender como o mundo subjetivo se constitui a partir da intersubjetividade. Nesse sentido, recusa qualquer distinção entre esquema conceitual e conteúdo empírico, colocando o sujeito em contato direto com a realidade empírica e social. Para evitar o Teatro Cartesiano, Davidson realiza alterações no argumento da tradução radical de Quine. Para Quine, a tradução radical se apoia nos estímulos sensoriais disponíveis ao tradutor durante o processo. Davidson considera isso uma forma de subjetivismo e passa a apoiar o processo de interpretação radical nos eventos externos ao sujeito. O subjetivismo é denominado por ele de “Mito do Subjetivo”, a ideia de “há misteriosas entidades ‘diante da mente’ que vêm entre nossos pensamentos sobre o mundo e o mundo em si, o que eu chamei de ‘intermediários epistêmicos’” (DAVIDSON, 1998/2005a, p. 321, aspas duplas do original substituídas por aspas simples, tradução nossa). Apesar disso, Davidson não deixa de reconhecer a existência de uma assimetria epistêmica entre os pontos de vista de primeira pessoa e de terceira pessoa¹³.

1.1.1. A mente descentralizada de Dennett

Tanto Dennett como Davidson se colocam contrários ao modelo do Teatro Cartesiano. Ambos entendem que uma reintegração do sujeito à realidade objetiva (após a ruptura) pode evitar esse problema. Os argumentos de Davidson serão mais desenvolvidos ao longo do segundo capítulo. Aqui se falará brevemente da proposta de Dennett de um modelo

¹³ Ver: seção 1.3.

alternativo, uma concepção do mental que evita a noção de Teatro Cartesiano. Dennett oferece o modelo dos rascunhos múltiplos e, posteriormente, o modelo da fama no cérebro¹⁴.

Muitas pessoas já tiveram a experiência de dirigir enquanto conversam com alguém ou pensam sobre algum assunto. Ao final do percurso, não se lembram mais do que aconteceu durante o caminho, das ultrapassagens que fizeram ou do sol que se pôs enquanto dirigiam. Isso, na concepção do autor, não significa que o motorista não estava ‘consciente’¹⁵ do que acontecia enquanto estava dirigindo, apenas que os eventos não se fixaram na memória. Isto é, a pessoa dirigia conscientemente e, por essa razão, foi capaz de executar os movimentos que lhe permitiram chegar viva ao seu destino (Dennett, 1991b, p.137-8, tradução nossa).

O autor argumenta que as nossas experiências não têm hora e local marcados para acontecerem, elas ocorrem em tempos e espaços mínimos. Além disso, a quantidade de informações é enorme. Porém, nossos cérebros foram projetados¹⁶ por um processo de seleção natural para lidar com o tempo e o espaço em escala macroscópicas. Isso levaria à adoção espontânea de um hábito quando o ser humano precisa olhar para dentro, e olhamos para o mundo interno como se ele fosse o mundo externo e os eventos ocorressem em escala macroscópica. O resultado é uma edição imediata das experiências que se tornam conscientes. Note-se, aqui, que se o sujeito precisar se remeter às suas experiências conscientes em outro momento, o processo editorial emitirá uma versão final diferente. Nesse sentido, o processo editorial é bastante contingente, assim como são contingentes os conteúdos que se tornam conscientes. As experiências se organizam por conta própria quando o sujeito precisa responder por elas.

Dennett (1991b) apresenta vários exemplos do modelo dos rascunhos múltiplos em ação no dia-a-dia das pessoas. O exemplo mais paradigmático desse processo editorial é o que ocorre quando assistimos um filme. O filme é apresentado como uma sequência de quadros estáticos em um intervalo muito pequeno (geralmente, 24 por segundo), o que gera a impressão de estarmos vendo movimentos contínuos. Se o filme for dublado, há pessoas que não reparam nas diferenças entre os movimentos faciais do ator e os sons proferidos. O processo de edição realizado pelo cérebro, na concepção de Dennett, conferiria às pessoas essa impressão de correspondência entre o que veem e escutam. Durante frações de segundo, o cérebro faria uma série de correções, adicionaria detalhes e faria alterações na ordem das

¹⁴ Sobre esses modelos, há um desenvolvimento em Fagundes (2009).

¹⁵ O termo, aqui, significa globalmente acessível para o comportamento ou outros estados mentais, não deve ser confundido com “consciência fenomênica” ou “*qualia*”. Adiante, se tratará dessa outra concepção de consciência.

¹⁶ Para Dennett (2017b, p. 53-54), uma importante “inversão” de Darwin na história do pensamento foi mostrar como é possível projetar máquinas incríveis sem saber como. É isso que o processo de seleção natural fez. Essa ideia já estava presente em textos anteriores do autor, como o “Consciousness Explained” (1991b).

experiências até produzir essa sensação de uma experiência contínua e integrada. O mesmo que acontece quando assistimos um filme, segundo o autor, aconteceria também quando observamos a realidade ao nosso redor.

Por exemplo, uma vez que sua cabeça se move pouco e seus olhos se movem muito, as imagens na sua retina flutuam constantemente, exatamente como as imagens dos filmes caseiros feitos por pessoas que não conseguem deixar a câmera sem sacudir. Mas não é assim que nos parece. As pessoas frequentemente ficam surpresas ao aprender que sob circunstâncias normais, seus olhos dardejaram rapidamente, em torno de cinco ligeiras fixações por segundo e que esse movimento, assim como o movimento de suas cabeças, é editado precocemente no processo do globo ocular até... a consciência (DENNETT, 1991b, p. 111, tradução nossa).

Outro exemplo interessante é o de duas lâmpadas que estão lado a lado, sendo uma vermelha e outra verde. Essas duas lâmpadas piscam alternadamente a um curto intervalo. Os sujeitos submetidos à experiência de olhar para isso relatam ter impressão de que a luz se movimenta de um lado para o outro mudando de cor ao longo do processo. Essa impressão ocorre a partir da primeira vez em que as lâmpadas piscam. Ou seja, para o sujeito, a impressão é que uma lâmpada vermelha caminha e se torna gradualmente verde, como se houvesse uma espécie de premonição. O modelo dos rascunhos múltiplos compreende essa situação como uma aparência de movimento e mudança de cor que é produzida retrospectivamente no cérebro. As experiências ocorreriam massivamente no cérebro, em uma temporalidade diferente daquela com a qual nossos cérebros estão acostumados (isto é, a temporalidade dos eventos com que nós e nossos ancestrais precisamos lidar no mundo exterior). A impressão que se produz é de um mundo interno que tem o tempo e o espaço do mundo externo.

Embora as experiências aconteçam de modo altamente fragmentado e desorganizado, elas passam por processos de interpretação. Esses processos podem vir do próprio sujeito da experiência, no momento em que procura narrar a alguém ou a si mesmo seus eventos internos ou podem vir de outro intérprete. Acontece, segundo Dennett, que quanto maior o distanciamento em relação aos eventos mentais, mais eles parecerão integrados. Quando mais forem vistos de perto, mais desorganizados e fragmentados. As narrativas que o sujeito conta sobre suas experiências exigem um alto grau de organização e interpretação, mas isso é feito naturalmente, resultado da necessidade que tivemos (e continuamos tendo) de atribuir estados mentais às outras pessoas.

O modelo da fama no cérebro foi elaborado posteriormente (DENNETT, 2005b). Não há nada de substancial que o diferencie do modelo dos rascunhos múltiplos, mas Dennett o supôs mais afável à imaginação, principalmente após o desenvolvimento de alguns estudos

experimentais na psicologia cognitiva, em especial os experimentos que mostram o alto grau de cegueira que temos às mudanças em situações cotidianas.

Simons e Levin (1997) fizeram alguns experimentos interessantes em relação a isso. Eles mostravam aos sujeitos pares de filmes em que mudavam apenas elementos específicos. Porém, em muitos casos, o elemento alterado era extremamente central, indo desde a roupa do protagonista até o próprio protagonista. Quando alteraram a roupa, incluindo sua cor, apenas 20% dos sujeitos foram capazes de perceber. Quando alteraram o próprio ator, o percentual aumentou para 33%. A maior parte dos sujeitos foi capaz de narrar a cena com detalhes, mas não percebeu a mudança. Por último, os experimentadores resolveram testar casos reais, mudando o ator em plena interação com o sujeito, e mesmo assim muitos não foram capazes de perceber a mudança. Esses experimentos são conhecidos como “cegueira à mudança”, mostram nosso alto grau de cegueira frente ao que está diante de nossos olhos.

O modelo da fama no cérebro considera que certos conteúdos, em determinados momentos, se tornam globalmente acessíveis para manifestação comportamental, são os conteúdos que se tornariam “famosos”. Isso ocorreria de modo totalmente contingente, sem nada para guiar esse processo e determinar quais conteúdos se tornarão acessíveis para manifestação comportamental. Assim, o fato de alguns sujeitos nos experimentos de Simons e Levin não terem percebido a mudança tem a ver, conforme o modelo da fama no cérebro, com aquele conteúdo não ter se tornado disponível para controle comportamental naquele momento para aqueles sujeitos.

Ao elaborar o modelo da fama no cérebro, Dennett parece estar mais preocupado em combater a noção de consciência fenomênica como algo privado e diretamente acessível ao sujeito do que com compreender a constituição do sujeito linguístico. Por isso, nesse livro (DENNETT, 2005b), não vemos uma ênfase no processo de interpretação das experiências próprias e alheias. De todo modo, Dennett continua, aqui, em busca de argumentos que permitam uma perspectiva alternativa ao Teatro Cartesiano e continua buscando mostrar as possíveis vantagens de uma abordagem do mental a partir do ponto de vista externo.

O modelo do Teatro Cartesiano ajuda a compreender o tipo de concepção de mental que Dennett busca combater. Para ele, não deve haver uma linha demarcando o sujeito do mundo, não deve haver um castelo particular do sujeito onde a introspecção reine. Não é que o sujeito não tenha estados mentais, não é que ele não perceba o mundo, nem que ele seja incapaz de qualquer tipo de introspecção. É que os pensamentos são construídos de fora para dentro a partir do processo de atribuição de estados mentais e, após construídos, mantém uma relação de troca constante com o mundo exterior.

1.2. Pausa para os *qualia*

A noção de estados mentais não disposicionais gera alguns problemas filosóficos bastante complicados, razão pela qual este trabalho procurará seguir por outro caminho a partir do segundo capítulo, tratando a mente como algo relacional. Deve-se reconhecer, porém, que os estados mentais considerados intrínsecos e inefáveis, os *qualia*, têm sido muitas vezes considerados o cerne do problema mente/corpo. Por isso, uma crítica comum ao interpretivismo é que a posição ignora os *qualia*. Argumentar-se-á que, dadas as confusões e aporias geradas pelos *qualia*, a busca de um outro tipo de abordagem do mental pode ser bastante apropriada.

Como disse Dennett (2013, p. 302, tradução nossa): “Uma tarde, sobre uma boa garrafa de Chambertin, o filósofo Wilfrid Sellars me disse ‘Dan, *qualia* são o que faz a vida valer a pena!’”. Vale considerar que Sellars é um autor bastante influente para o interpretivismo, como se tentará mostrar adiante. Outros autores, como Block (1990), Nagel (1974/2005) e Tye (2017), também têm se dedicado a mostrar que esses estados mentais não podem ser deixados de lado por quem trata do mental. Nesse sentido, a questão vale uma breve parada para explicar por que este trabalho se desviará da questão dos *qualia*.

Será tomada uma versão restrita do termo *qualia* como estados mentais não relacionais que reúnem um conjunto específico de características. Porém, essas características costumam estar presentes quando se fala de *qualia*. Por isso, embora restrita, é uma versão que se tornou muito aceita e reconhecida. Se eles forem tomados como estados mentais capazes de participar de uma cadeia causal com consequências comportamentais, os argumentos aqui apresentados não caberão. Assim, embora o trabalho que se segue a este capítulo tenha uma ênfase nas atitudes proposicionais, estados mentais linguísticos, ele não vai deixar de considerar a existência de estados mentais não linguísticos, desde que disposicionais. Esses estados precisam ser levados em conta, sim, pois a mente foi construída a partir de um processo evolutivo no qual a linguagem só surge em estágios muito tardios. Assim, para que ela pudesse surgir, foi preciso uma base de estados mentais pré-linguísticos que lhe desse sustentação. Então, o problema não é reconhecer os estados mentais que não são proposicionais, mas sim ter que lidar com supostos estados mentais totalmente não relacionais.

Em um artigo de 1988, Dennett propõe quatro condições para tentar unificar o conceito a que se referem os defensores dos *qualia*. De acordo com ele, o termo “*qualia*” se referiria a supostos estados mentais simultaneamente *inefáveis*, *intrínsecos*, *privados* e

diretamente apreensíveis pelo sujeito. Em outras palavras, eles só aparecem por meio da introspecção; são indescritíveis; não são relacionais (“a cor vermelha pode ser provocadora de ansiedade para algumas pessoas, mas essa disposição subjetiva não é um *quale* de vermelho” (DENNETT, 2013, p. 298, tradução nossa)) e aparecem apenas para o sujeito, não podem ser compartilhados; estão presentes na consciência e são o que há de mais íntimo. Essa reunião de características é o que provoca determinadas aporias, evitáveis a partir de uma compreensão relacional dos *qualia*.

Muitos autores com uma tendência dualista afirmam que a abordagem interpretivista trata do mental de forma apenas parcial e que o verdadeiro problema permanece sem explicação. Davidson, por exemplo, se restringe apenas às atitudes proposicionais. Porém, argumentam alguns autores (CHALMERS, 1996, por exemplo) que o mental é fundamentalmente constituído por estados fenomênicos, isto é, as características qualitativas e subjetivas das nossas sensações. De acordo com essa concepção, tais estados não podem ser compartilhados, são privados e, conseqüentemente, não podem ser expressos em termos de atitudes proposicionais. Por mais que falemos deles, jamais poderíamos levar o nosso interlocutor a uma compreensão completa do que sejam os nossos estados subjetivos internos.

Chalmers (1996, p. 377) considera a abordagem de Davidson interessante por negar a possibilidade de leis psicofísicas estritas. Porém, defende que o problema realmente relevante e difícil acerca do mental [“the hard problem”] não é uma explicação acerca de como as atitudes proposicionais possam causar o comportamento. O difícil, segundo ele, é explicar como pode haver uma vida subjetiva interna, fenomênica, acompanhando os eventos no mundo físico. Essa vida subjetiva interna, na concepção de Chalmers, seria composta de experiências fenomênicas, justamente aquilo com o que os filósofos da mente deveriam se ocupar. Ela constituiria o verdadeiro problema do mental e não poderia ser expressa por meio de estados mentais disposicionais. Para defender a sua posição, Chalmers precisa apelar para um tipo de epifenomenalismo¹⁷. Ele defende que a cadeia causal do mundo natural seria a mesma independentemente dos *qualia*¹⁸. Chalmers não consegue abandonar o Teatro

¹⁷ “Epifenomenalismo” é posição segundo a qual os estados mentais são subprodutos dos estados físicos, mas não possuem qualquer poder causal sobre eles. Segundo essa posição, os nossos estados físicos causam estados mentais, mas estes não provocam nenhuma influência sobre aqueles, nem há relações causais entre os próprios estados mentais. Ver: seção 1.3.2.

¹⁸ Em um artigo mais recente, Chalmers (2018, p. 17, tradução nossa) afirma que a concepção dos *qualia* é amplamente rejeitada atualmente e que “é muito mais comum sustentar que as experiências são (ou parecem ser) estados relacionais”. Até aqui, essa nova perspectiva traz alívio, pois parece que as complicações provocadas pela noção de *qualia* foram finalmente deixadas para trás, mas o alívio é apenas momentâneo. Chalmers continua: “Por exemplo, a experiência de esverdeamento [“greenness”] não envolve uma qualidade ‘verde’ simples, em vez disso parece envolver reconhecimento [“awareness”] do esverdeamento, a cor.” Parece que, mesmo que não fale de *qualia*, Chalmers continua defendendo que a experiência tem um caráter não-relacional,

Cartesiano. O sujeito, para ele, precisa ter um ambiente íntimo inacessível a qualquer outra pessoa: o lugar onde se dão as experiências subjetivas.

Muitas vezes associados às experiências sensoriais, mas também associados às emoções, ao prazer e à dor - *qualia* seriam o que constitui as nossas experiências subjetivas, algo que só pode ser vivido pelo sujeito, que não se compartilha, só meu. A ideia dos *qualia* como constituinte básico da vida mental de um sujeito tem sido defendida ao longo da história recente da filosofia da mente a partir de uma série de experimentos de pensamento filosóficos. Serão abordados aqui dois deles: o argumento do zumbi e o do espectro invertido.

Afirmam os defensores dos *qualia* que não há como falar muita coisa sobre os *qualia*, uma vez que eles são subjetivos e a fala é intersubjetiva, apenas se pode postular que cada um se volte para dentro no intuito de “comprovar sua existência”. Apesar dessa dificuldade em relação ao discurso, são considerados evidentes demais para que se possa negá-los. Porém, será que eles não poderiam ser estados disposicionais que participassem de uma corrente causal com manifestação comportamental? A resposta de muitos autores é que não. O argumento do zumbi é um importante experimento de pensamento usado para convencer as pessoas do quanto os *qualia* são intratáveis a partir do discurso. Zumbis seriam seres imaginários que se comportariam exatamente como sujeitos de estados mentais, mas que careceriam totalmente de *qualia*.

Uma desvantagem desse tipo de concepção do mental é que torna impossível um estudo objetivo do mental. Uma vez que os *qualia* não teriam qualquer expressão comportamental, não seria possível falar sobre eles. Assim, quando alguém dissesse que adora chocolate, a experiência subjetiva que tem do sabor do chocolate não poderia ter qualquer relação com aquela frase, nem com a crença da pessoa de que adora chocolate. Para Dennett (1995), o argumento provém de um mau uso da imaginação. Quando o argumento é cuidadosamente considerado e suas consequências são levadas a sério, ou ele nos leva a paradoxos difíceis de serem superados ou ao eliminativismo de *qualia*.

Quiçá se tema que os *qualia* sejam o último reduto do sujeito, de modo que deixar os *qualia* de lado estilhaçaria os limites da subjetividade. Porém, a compreensão do sujeito como algo que não está isolado do mundo objetivo não precisa negar a autoridade de primeira pessoa acerca dos estados mentais. Davidson, por exemplo, deliberadamente se abstém de tratar dos *qualia*, mas nem por isso abre mão da autoridade de primeira. A noção de sujeito

as propriedades fenomênicas, mas é acompanhada por um reconhecimento dessa própria experiência, algo que fará parte da cadeia causal. Mas o problema é justamente esse: postular estados mentais não relacionais. O que causaria, então, o reconhecimento?

desempenha um papel central em seu pensamento, enquanto o subjetivismo é rechaçado. Talvez não seja preciso negar as experiências sensoriais, nem que sejam prazerosas ou dolorosas, desde que não precisemos isolar essas experiências dos demais estados mentais.

1.2.1. *As cores e as outras mentes*

Locke (1690/2012) denomina “ideia” aquilo que está na mente, e não nos objetos, embora os objetos tenham a potência de provocar ideias nas mentes. Para ele, há dois tipos de qualidades nos objetos que percebemos: as qualidades primárias, que estariam nos objetos independentemente dos nossos sentidos (extensão, figura, número e movimento), e as qualidades secundárias, isto é, as que se apresentam aos nossos sentidos e estão nos objetos apenas enquanto potências capazes de nos provocar tais experiências (cores, gostos, cheiros etc.). As qualidades secundárias estariam restritas à mente do sujeito. Por isso, na concepção de Locke, não é possível ao sujeito saber com certeza o que se passa na mente de outro sujeito.

Nossas *ideias* simples não poderiam ser ditas falsas *ainda* que a diferente estruturação dos órgãos de cada homem os dispusesse de tal maneira que *um mesmo objeto produzisse numa mente ideias diferentes das que produz noutra*, como se a *ideia* produzida pela *violeta* na mente de um homem fosse a mesma produzida pelo *girassol* na mente de um outro, e *vice-versa*. Seria impossível saber que isso acontece, pois não é possível instalar a mente de um homem no corpo de um outro, para perceber quais aparições são produzidas por aqueles órgãos: não se confundem nem *ideias*, nem seus nomes, que não são, nenhum deles, *falsos*¹⁹ (LOCKE, 1690/2012, p. 416).

No trecho acima, encontramos o argumento segundo o qual se uma pessoa tiver experiências subjetivas de cores radicalmente diferentes das minhas (ou das de outra pessoa qualquer), não teremos como detectar essa diferença, pois a experiência é absolutamente restrita ao sujeito, ao passo que a linguagem é compartilhada. Esse argumento é denominado “argumento do espectro invertido”²⁰.

O argumento apresenta versões contemporâneas também (por exemplo, Block, 1990). A ideia é que, se uma pessoa tem a experiência subjetiva de verde diante de um tomate maduro, ela vai chamar aquele tomate de vermelho, uma vez que aprendeu publicamente o

¹⁹ Livro II, cap. XXXII. § 15.

²⁰ Atualmente, há pesquisas em psicologia mostrando que temos uma série de reações comportamentais diante das experiências de cores. Considera-se, portanto, que a inversão do espectro sem expressão comportamental é muito improvável (Byrne, 2006). Ainda assim, tais argumentos de base científica acabam não sendo acolhidos pelos filósofos que tendem a separar a experiência subjetiva de qualquer expressão comportamental. A razão para esse não acolhimento é que muitos argumentos filosóficos consideram não possibilidades materiais, mas sim lógicas.

vocabulário relativo às cores. Na versão de Block, algumas adaptações são feitas, já que o desenvolvimento das pesquisas em psicologia cognitiva mostrou que o espectro de cores é assimétrico, o que inviabilizaria essa possibilidade de inversão no mundo material. Embora o argumento apresente diferentes versões, não as adentraremos, pois giram em torno do mesmo eixo²¹.

Esse argumento serve para iluminar a ideia de que determinados estados mentais, a saber, os *qualia*, são acessíveis apenas para o sujeito. Vale notar que, se é possível as outras pessoas possuírem experiências subjetivas diferentes das minhas sem que eu tenha condições de perceber, talvez também possam simplesmente não possuir quaisquer experiências, e eu também não seria capaz de perceber. Nesse caso, então, a possibilidade filosófica de estarmos interagindo com zumbis estaria aberta²². Se seguimos por essa linha de pensamento, vamos inevitavelmente desembocar no ceticismo acerca de outras mentes. Há como evitá-lo?

1.2.2. Zumbis: uma concepção de qualia a ser revista

A concepção de *qualia* que pretendemos evitar é aquela segundo a qual *há algo para o sujeito que é como estar em um determinado estado mental*. Essa concepção, tornada célebre no artigo “Como é ser um morcego” de Nagel (1974/2005), parece carregar um comprometimento ontológico com algo que se manifesta internamente para o sujeito. De acordo com essa visão, os *qualia* seriam o aspecto fundamental da vida consciente e estariam sempre disponíveis para o sujeito, mas só para ele, criando um abismo entre dois aspectos da realidade: o aspecto objetivo e o aspecto subjetivo. É justamente essa ideia de entidade isolada dos demais estados mentais e cerebrais que carrega grandes dificuldades. Dennett (2013, p. 296-297) considera um erro a ideia de que os *qualia* sejam estados mentais isolados. As nossas experiências sensoriais se combinam com outros estados mentais, como desejos crenças e comportamentos diversos, verbais ou não (nesse sentido, o argumento também pode se estender aos animais não-humanos, ainda que desprovidos de linguagem simbólica). Ao compreendê-los de forma relacional, não sobra nada encarcerado na mente do sujeito.

²¹ Para um aprofundamento, ver Byrne (2006).

²² Para que isso não soe insólito demais, deve-se lembrar novamente que os experimentos de pensamento em filosofia geralmente não tratam de possibilidades materiais. Não é que a minha banca de defesa possa realmente estar lendo a tese de uma zumbi. Os experimentos de pensamento lidam com a imaginação e, muitas vezes, consideram o imaginável como metafisicamente possível, ainda que não ocorra em nosso mundo. Falando de modo bastante geral, se não houver uma contradição interna entre as proposições daquele experimento de pensamento, ele é metafisicamente possível, ainda que não ocorra neste mundo em que estamos.

O argumento do zumbi é comumente utilizado para defender essa concepção de *qualia*: sua existência como algo resistente a qualquer abordagem de um ponto de vista objetivo. De acordo com Chalmers (1996, p. 94, tradução nossa), por exemplo, é possível a existência de “alguém ou algo fisicamente idêntico a mim (ou a qualquer outro ser consciente), mas completamente destituído de experiências conscientes”. Vale a pena, aqui, explicar que o zumbi não é simplesmente uma duplicata molecular de um ser consciente da pele para dentro. Ele mantém com o mundo físico as mesmas relações que um ser consciente e apresenta os mesmos comportamentos, inclusive os verbais. Yudkowsky define da seguinte maneira:

Nosso “zumbi”, no uso filosófico do termo, é putativamente um ser que é exatamente como você em *todos* os aspectos - comportamento idêntico, fala idêntica, cérebro idêntico; cada átomo e quark *exatamente* na mesma posição, movendo-se de acordo com as mesmas leis causais de movimento - *exceto* que o nosso zumbi não é consciente (YUDKOWSKY, 2015, p. 1012, aspas internas e grifos no original, tradução nossa).

Ocorre que, para poder dizer as mesmas coisas e exibir os mesmos comportamentos em geral, presume-se que o zumbi possua as mesmas atitudes proposicionais da sua contraparte consciente. Afinal, para que um ser possa ser considerado um zumbi, ele precisa ser capaz de falar sinceramente acerca de suas crenças, tal qual um ser consciente. Suponha que eu tenha uma zumbi gêmea. Ela é minha substituta em um mundo no qual eu não existo. Lá, ela faz tudo o que eu faço aqui. Então, se eu digo que neste momento tenho plena confiança nas instituições jurídicas brasileiras, ela diz a mesma coisa. Se formos ambas submetidas a um teste no detector de mentiras, seremos igualmente desmascaradas. Em tudo, nossos corpos se manifestam exatamente da mesma maneira.

O que nos provoca tais reações perante o detector de mentiras não é dizermos algo em que não acreditamos? Ora, mentir é justamente afirmar algo que se crê como falso. O comportamento de mentir está sempre ligado a uma crença. Então, a minha contraparte zumbi pensa como eu, possui as mesmas crenças, as mesmas dúvidas, os mesmos desejos e gostos que eu. Mas nós temos uma enorme diferença em relação uma à outra: eu possuo *qualia* e ela não possui. Ao entrar na sala do detector de mentiras, eu fico ansiosa, sinto as mãos frias e minha testa fica úmida. A testa dela também fica úmida, as mãos gelam, mas ela não tem a sensação subjetiva do frio, embora ela creia que tem a sensação e diga sinceramente que está sentindo as mãos frias. Ao ver uma barata andando pela sala, o estômago dela revira, os pelos se arrepiam e, assim como eu, ela afirma sinceramente que está sentindo muito nojo, mas com a diferença que ela não tem a experiência subjetiva do nojo.

Um aspecto curioso dessa diferença tão radical entre mim e ela é que ninguém, absolutamente ninguém, é capaz de notá-la. Qualquer pessoa que observe nossos comportamentos tenderá a atribuir a ambas a mesma vida subjetiva. Se um dia eu e ela pudéssemos marcar um encontro, eu não teria meios para saber que ela carece de experiências e ela não poderia saber que eu tenho uma vida interna rica em experiências que lhe carecem. Isso é um aspecto importante do argumento. Sua força reside no fato de que nós nunca podemos afirmar com certeza o que se passa na vida subjetiva alheia.

Contudo, também é justamente isso o que conduz o argumento a uma aporia. Assim como ela não tem como saber que eu tenho uma vida subjetiva abundante, também não tem como saber que sua vida subjetiva é vazia. Se o saber passa pela crença e minha gêmea zumbi crê que possui experiências tão variadas e abundantes como as minhas, então ela não sabe de sua ausência de experiências²³.

Ora, mas que coisa é essa que chamamos de *qualia*, que cremos que possuímos, que uma suposta contraparte zumbi nossa também crê que possui, mas nós estamos certos e ela está errada em sua crença? Diante de minhas limitações epistêmicas, nada me poderia garantir que eu mesma não seja um zumbi. Essa conclusão não parece ser o que os defensores do argumento do zumbi desejariam defender. Eles querem defender a existência de estados subjetivos diretamente apreensíveis pelo sujeito consciente e absolutamente internos, sem expressão comportamental. Porém, o que eles conseguem é nos levar à conclusão de que não sabemos se somos ou não somos zumbis. Nesse sentido, o argumento do zumbi, comumente usado para defender os *qualia*, talvez fosse melhor orientado se utilizado para rejeitar a noção de *qualia*.

O argumento do zumbi muitas vezes é apresentado em termos de mundos possíveis. Assim, de acordo com essa versão do argumento do zumbi, seria possível (não que isso ocorra na prática, mas seria uma possibilidade teórica) um mundo exatamente como o nosso, com a única diferença de que, nesse mundo, as pessoas e os animais *não* teriam experiências subjetivas internas. Sendo assim, seria possível concluir que essas experiências deste mundo são algo além de todos os aspectos físicos do mundo, uma espécie de “bônus”, um acréscimo sobrenatural ao mundo natural, mas um acréscimo que não é capaz de provocar nenhum efeito sobre a realidade física. Tudo permaneceria exatamente como está.

²³ Chalmers diria que minha crença é justificada pela minha experiência fenomênica. “O que é que justifica nossas crenças sobre nossas experiências [...]? Eu penso que a resposta a isso é clara: é ter as experiências que justifica as crenças” (CHALMERS, 1996, p. 196, tradução nossa). Eu *saberia* das minhas experiências, em contraste com a mera opinião infundada da minha zumbi. Porém, o que seria a causa da justificação da minha crença? Não pode ser a experiência fenomênica, já que ela não possui poderes causais. O argumento, nesse ponto, parece não ter mais saída.

Vale ressaltar que, no mundo zumbi, haveria muitos artigos e livros de filosofia falando sobre zumbis e sobre as experiências subjetivas como epifenômenos, de modo que a própria experiência consciente não poderia ter qualquer relação com o impulso filosófico de escrever sobre ela. Nesse sentido, Yudkowsky (2015) argumenta que o argumento leva à conclusão de que, se um filósofo escrever algo sobre a consciência e isso for verdadeiro, será verdadeiro por pura coincidência. Nenhuma crença sobre experiências conscientes pode ser causada pela própria experiência, já que o mundo seria o mesmo, havendo ou não essas experiências. Os filósofos escreveriam sobre elas de qualquer forma. Portanto, se as crenças sobre experiências carecem de justificação no mundo zumbi, também carecem no nosso mundo. Seria um milagre contingente do nosso mundo se as crenças acerca das experiências epifenômicas fossem corretas.

O argumento do zumbi é um exemplo de um problema maior. O ceticismo oscila entre as outras mentes e a própria mente como resultado de uma concepção que isola os estados mentais do mundo objetivo. Esse tipo de conclusão pode ser evitado, desde que se mudem as premissas. A mente não se constitui de um teatro frente ao qual uma audiência subjetiva possua uma série de experiências. Ao contrário, se constitui gradualmente no processo de interação do organismo com o mundo objetivo e os outros organismos.

1.3. A minha mente e as outras mentes

Quando falamos de nossos próprios pensamentos, temos um grau bem alto de convicção, mas quando atribuímos pensamentos aos outros, a dúvida e a possibilidade do erro se fazem presentes. Surgem aqui duas questões filosóficas conectadas: a autoridade de primeira pessoa acerca dos próprios estados mentais e o problema do ceticismo acerca das outras mentes. Davidson (1987/2001a) pretende enfrentar ambas, mostrando que sua concepção do mental não ameaça o reconhecimento de uma assimetria entre o conhecimento de primeira pessoa acerca dos próprios estados mentais e o conhecimento que temos acerca dos estados mentais alheios. Além disso, seu tipo de abordagem fecha as portas de entrada para o cético acerca das outras mentes.

A conexão entre o problema da autoridade de primeira pessoa e o problema tradicional das outras mentes é evidente, mas como levanto o primeiro problema, existem duas diferenças importantes. A autoridade de primeira pessoa é o problema mais restrito, uma vez que o tratarei apenas na medida em que se aplica a atitudes proposicionais como a crença, o desejo e a intenção; estar satisfeito, admirado, assustado ou orgulhoso de que algo seja o caso. Mas não analisarei o que frequentemente se considera que é central no problema das outras mentes: as dores e outras sensações e o conhecimento, a memória, a atenção e a percepção tanto

dirigidas a objetos como a pessoas, ruas, cidades, cometas e outras entidades não proposicionais. *Parece-me que o que se mantém para as atitudes proposicionais há de ser relevante para as sensações e o restante, mas não vou explorar aqui essa conexão* (DAVIDSON, 1984/2001g, p. 25-6, grifo nosso, tradução nossa).

O problema das outras mentes se faz presente quando uma perspectiva subjetivista é adotada. Porém, diante de uma abordagem externista, talvez ele não se apresente. De modo bem geral, pode-se dizer que externismo acerca do mental é a posição segundo a qual o significado ou o conteúdo de certos estados mentais é ao menos parcialmente determinado pelo ambiente externo²⁴. De acordo com Davidson (1987/2001a), alguns autores (como Putnam, Burge e ele próprio) defendem que os estados mentais proposicionais, por possuírem conteúdo linguístico, só podem ser formados por meio de relações com o ambiente externo, inclusive o ambiente social. Ora, se aceitamos que tais estados mentais dependem de algum tipo de relação externa com o ambiente socialmente compartilhado com outros sujeitos, a perspectiva cética acerca das outras mentes não se coloca. Contudo - afirma Davidson - Putnam e Burge fazem uma diferenciação entre estados mentais intrínsecos e relacionais. Certos estados mentais, para eles, ainda pertenceriam exclusivamente ao sujeito e a autoridade de primeira pessoa se conservaria apenas em relação a tais estados mentais. Haveria, então, na concepção desses autores, estados mentais completamente isolados do mundo externo, resistentes a qualquer tipo de manifestação comportamental.

Esse tipo de externismo que separa os estados mentais entre intrínsecos e relacionais dá origem ao seguinte problema epistemológico: se o conteúdo de meus estados mentais proposicionais depende, ainda que parcialmente, de fatores externos ao sujeito, então como poderia o sujeito ter certeza acerca dos conteúdos de seus pensamentos? Um exemplo de Parent (2013): se para Descartes saber que está pensando sobre carvalhos ele precisa saber o que são carvalhos no mundo exterior, então ele não tem como determinar o conteúdo de seus próprios pensamentos apenas sentado em sua poltrona.

O problema de conciliar essa versão de externismo com o autoconhecimento acerca dos estados mentais é que surge uma lacuna entre os estados mentais intrínsecos e os relacionais e essa lacuna leva a uma impossibilidade de dar sentido às experiências intrínsecas. Se elas não mantêm contato com os estados mentais relacionais, então, ao olhar para seu mundo interior, o sujeito encontraria apenas experiências confusas. Como poderia conectar a sua experiência interna de vermelhidão ao termo “vermelho”, sendo que o conteúdo do termo depende causalmente das relações entre o sujeito e o objeto vermelho no mundo?

²⁴ Ver: nota de rodapé 4.

A tradição cartesiana é internista. A atribuição a um sujeito de uma propriedade psicológica não depende de nada que seja externo a esse sujeito. Mas isso dá origem ao ceticismo acerca das outras mentes. Ora, se duvidamos de que os outros possuam mentes, toda nossa comunicação e vida social também ficam em cheque e, no fim das contas, talvez eu possa ser simplesmente um *cérebro numa cuba* dentro do laboratório de um cientista maligno que me controla por meio de eletrodos e me faz crer que sou uma doutoranda diante de um computador escrevendo justamente sobre isso (PUTNAM, 1981). Talvez nada mais exista além da minha própria mente.

Em uma perspectiva externista, por outro lado, o problema das outras mentes pode ser deixado de lado, pois os estados mentais se constituem parcialmente a partir de suas relações causais com o mundo externo. Contudo, um novo embaraço filosófico pode surgir em algumas versões de externismo: o problema de se a pessoa sabe em que estado mental ela própria está. Enquanto se apaga o problema das outras mentes, talvez o problema da minha própria mente se acenda. Se sim, ocorre um deslocamento do ceticismo para um lado ainda mais contra intuitivo. No fundo, esses dois problemas têm a mesma origem: a tradição de ruptura entre os mundos objetivo e subjetivo.

Davidson desenvolve seu argumento em torno dos estados mentais proposicionais e não se compromete com a existência de estados mentais que sejam isolados e restritos ao sujeito. Ao contrário, como podemos notar na última frase da citação acima, ele pensa que suas conclusões podem iluminar tais estados comumente considerados intrínsecos, mas não se ocupará com o desenvolvimento desse ponto. O que ele pretende mostrar é que a sua versão do externismo acerca dos estados mentais proposicionais não compromete a defesa da ideia de que o sujeito possui uma autoridade especial acerca daquilo que ele mesmo crê ou pensa.

Para Davidson (1988/2001j), a melhor maneira de compreendermos essa conciliação é a partir dos fatos por meio dos quais aprendemos a linguagem e interpretamos as palavras. Essa aprendizagem provém dos objetos e circunstâncias que cercam o aprendiz. Há uma relação de triangulação entre aprendiz, instrutor e mundo. Vale notar que a triangulação se faz presente sempre que há duas criaturas em interação entre si e com o mundo exterior. Portanto, a triangulação é anterior à linguagem e serve de suporte para que ela possa surgir²⁵.

Outro ponto que merece ser ressaltado é que, após o surgimento da linguagem convencional, o pleno funcionamento do processo de triangulação depende de que os

²⁵ Note-se também que o ceticismo acerca das outras mentes é elaborado no contexto da linguagem. Portanto, se o argumento da triangulação é correto, o ceticismo só pode ser elaborado após a aprendizagem pública da linguagem (SILVA FILHO, 2002, p. 161).

envolvidos no diálogo atribuam crenças majoritariamente verdadeiras uns aos outros. Cada um precisará partir do pressuposto que os sinais linguísticos emitidos pelo outro possuem uma conexão apropriada com o mundo. Apenas assim é que a comunicação linguística se torna possível. Além disso, se conseguimos algum sucesso ao nos comunicarmos, é porque de fato o que se passa na mente de nosso interlocutor coincide em grande medida com o que pensamos que se passa lá²⁶.

A perspectiva do intérprete que atribui crenças ao outro para poder compreender o que ele diz, assim, é em grande medida confiável. Ainda que o intérprete atribua muitas crenças falsas ao seu interlocutor, só poderá atribuí-las sobre um pano de fundo de crenças majoritariamente verdadeiras. Por isso, dentro do conjunto total de crenças do usuário de uma língua, a grande maioria delas deverá ser verdadeira. Caso contrário, a comunicação não seria possível. De modo bem geral, é a isso que se denomina “princípio de caridade”, um pressuposto para a interpretação bastante desenvolvido no pensamento de Davidson²⁷.

Diante dessa perspectiva externista de Davidson, portanto, sua estratégia para salvaguardar a autoridade de primeira pessoa acerca dos estados mentais consiste em apresentar o mental como sendo relacional. Deve-se abandonar a ideia de cisão entre o sujeito e o mundo. “A solução, no caso dos estados mentais (...) se trata de desfazer-se da metáfora dos objetos ante à mente” (1987/2001a, p. 68, tradução nossa).

Uma vez que tenhamos abandonado o mito do subjetivo – a ideia de que os pensamentos requerem objetos mentais – é natural pensar que a autoridade de primeira pessoa, o caráter social da linguagem e os determinantes externos do pensamento e o significado se deem conjuntamente. (DAVIDSON, 1987/2001a, p. 43, tradução nossa)

A proposta é que se abandonem as concepções de atitudes proposicionais como coisas que se apresentam perante o teatro mental. Se houvesse objetos na mente que fossem isolados do mundo externo, não poderíamos atribuir estados mentais a outras pessoas, não teríamos uma base para fazer isso. Ainda que haja estados mentais que não estejam diretamente conectados com o mundo exterior, eles estão conectados indiretamente, via outros estados mentais. Não é que Davidson defenda uma completa impossibilidade de erro do sujeito perante seus estados mentais. É possível haver erro, sim. Porém, a possibilidade do erro só é possível sobre um fundo de acertos.

Considerando a proposta de Davidson de que as crenças atribuídas a alguém por seu intérprete sejam majoritariamente verdadeiras, podemos nos perguntar sobre a possibilidade

²⁶ Ver: seção 2.3.

²⁷ Ver: seção 2.2.

de os *qualia* serem vinculados a essa cadeia causal. Nesse caso, certamente, teríamos que adotar uma concepção de *qualia* bastante diferente daquela apresentada acima e comumente adotada pelos defensores dos *qualia*. Porém, dadas as aporias epistêmicas a que fomos conduzidos pela concepção de estados qualitativos como inefáveis, privados, não relacionais e diretamente apreensíveis pelo sujeito, talvez uma reformulação do conceito seja bem-vinda.

De todo modo, não é necessário ignorar ou negar a existência de estados mentais não proposicionais, desde que não sejam isolados dos outros tipos de estados mentais. No caso humano, nossas experiências provavelmente têm potencial para ajustar nossos comportamentos, contribuir na constituição de nossas crenças e desejos e se fazerem presentes em nossas vidas sociais. As nossas crenças não estão isoladas de nossas experiências, e é por isso que temos uma autoridade de primeira pessoa acerca dos nossos estados mentais e podemos falar sobre eles.

Se digo sinceramente a alguém que estou sentindo calor, a fala se constitui a partir do que aprendi acerca do significado da palavra “calor”. Aprendi isso por meio de relações com as pessoas que me ensinaram a linguagem e com o ambiente externo. Então, a sensação que meu interlocutor me atribui também deriva de suas relações com o mundo, a linguagem e suas próprias experiências internas sendo, portanto, se não totalmente correta, suficientemente próxima daquilo que eu realmente sinto. Caso contrário, não seria possível me compreender nem se fazer compreendido por mim. As interações entre o sujeito, o mundo e os outros sujeitos tendem, conforme Davidson, a nos conduzir a crenças majoritariamente verdadeiras sobre o mundo, sobre minha mente e sobre as outras mentes. Os diversos estados mentais estão relacionados entre si, influenciam o comportamento e, no caso humano, constituem a linguagem e são constituídos por ela. Nesse sentido, uma abordagem relacional dos estados mentais é capaz de se desviar do ceticismo acerca das outras mentes sem descambar para o ceticismo acerca da própria mente.

1.3.1. Digressão sobre ontologia

Aqui, uma breve parada para tratar da noção de “ontologia”. Leclerc (2006) escreve um artigo no qual defende uma abordagem contextualista do predicado de existência, após mostrar por meio de exemplos que tal predicado de fato assume sentidos diferentes em contextos diversos. Essas diferenças orbitam um núcleo difuso, mas suficientemente denso para que possamos considerar todas como sentidos contextuais do mesmo termo. Em outras palavras, as diferentes noções contextuais não se constituem como ambiguidades. O mesmo

vale para o termo “ontologia”, sua compreensão também é contextual. Diante disso, resta tentar oferecer uma noção - ainda que pouco precisa - de “ontologia” no contexto desta tese.

Em seu último livro, Dennett parece associar a noção de “ontologia” a crença, ou a outros estados cognitivos dos animais e mesmo das máquinas:

“Ontologia” vem da palavra grega para coisa. Em filosofia, refere-se a um conjunto de “coisas” que uma pessoa acredita que existem, ou o conjunto de coisas definidas ou pressupostas por alguma teoria. O que está em sua ontologia? Você acredita em fantasmas? Então os fantasmas estão em sua ontologia juntamente com mesas, e cadeiras, e músicas, e férias, e neve e todo o resto. Provou-se mais do que conveniente estender o termo “ontologia” além do seu significado primário e usá-lo para o conjunto de “coisas” que um animal pode reconhecer e comportar-se apropriadamente com respeito a [elas] (independente de se é possível ou não dizer apropriadamente que os animais têm crenças) e - mais recentemente - o conjunto de “coisas” com as quais o computador deve ser capaz de lidar para fazer o seu trabalho (independente de se é possível ou não dizer apropriadamente que ele [“it”] tem crenças) (DENNETT, 2017, p. 60, aspas e grifos internos do original, tradução nossa).

Compreender “ontologia” como algo relativo meramente ao que alguém acredita, no contexto desta tese, contudo, não parece a melhor escolha. Afinal, a noção de “ontologia” aqui tem a ver com o conjunto de coisas que existem e que podem manter relações de causa e efeito, independentemente das crenças individuais das pessoas. Ainda que as crenças façam parte dessa mesma ontologia, elas não criam essa ontologia. Certamente, as coisas nas quais as pessoas acreditam, inclusive os fantasmas, devem participar da ontologia, uma vez que elas possuem influência causal sobre o comportamento. Porém, não há razão para simplesmente deixarmos algo de fora simplesmente porque ninguém acredita naquilo. A concepção de ontologia de Dennett não se restringe às crenças individuais de alguém, ela promove um elo entre sujeito e mundo, ainda que imperfeito e altamente contingente.

Isso fica mais claro quando Dennett amplia a noção para o reino animal ou mesmo para os computadores, aproxima-se mais do que aqui se quer significar por “ontologia”. Há coisas que um animal pode reconhecer e que, de alguma forma, direcionam o comportamento daquele animal (mas não são apenas isso, há outras coisas também). Essas coisas estão no mundo circundante do qual nós, os animais e os computadores fazemos parte. Nesse ponto, a noção oferecida por Dennett cabe bem ao nosso contexto, na medida em que não separa os ambientes interno e externo de um organismo ou máquina e admite as interações entre eles.

Aqui, não parece interessante pensar na minha ontologia, composta das coisas que eu acredito, a sua ontologia, composta das coisas que você acredita, a ontologia da minha gata, composta das coisas que influenciam em seus comportamentos, a ontologia do meu notebook... Há algo mais amplo, composto de todas essas coisas que influenciam ou não

nossos comportamentos individuais e até mesmo coisas nas quais não cremos ou que porventura não possamos conhecer diante de nossas limitações cognitivas, mas que estão aí no mundo, fazendo alguma diferença em como as coisas são, por mais que isso tudo seja altamente contingente. Podemos, então, falar neste contexto de uma realidade que envolve todas as coisas capazes de se relacionar, um mundo do qual nós participamos (coisas com as quais mantemos relações afetivas, coisas que criamos, coisas que modificam nossas vidas individual e coletivamente...), mas do qual também podem participar coisas das quais não fazemos a mínima ideia²⁸. Porém, há aspectos da realidade que podemos alcançar, enquanto outros estão além do nosso alcance.

A proposta de Dennett é muito interessante por relacionar os seres vivos com o ambiente a partir dos interesses comportamentais desses indivíduos. Sua proposta não leva ao completo relativismo ontológico, pois ele considera que há uma taxa alta de coincidência entre as ontologias das pessoas. “Alguns acreditam em bruxas e alguns acreditam em elétrons e alguns acreditam em ressonâncias mórficas. Mas há um núcleo enorme de ontologia que é compartilhado por todos os seres humanos a partir de uma idade muito precoce - seis anos de idade irão capturar quase tudo isso” (DENNETT, 2017b, p. 61, tradução nossa). A ontologia, na concepção dele, parece provir da epistemologia. Porém, neste momento convém termos uma noção, ainda que difusa, que ontologia e epistemologia não são o mesmo, ainda que as noções acabem por se tornar muito próximas no decorrer deste estudo.

Aqui, cabe chamar a atenção para um ponto importante: buscar temporariamente estabelecer que ontologia e epistemologia são coisas distintas não é uma ideia que precisa vir colada à ideia de que a ontologia seja algo fixo, imutável, universal ou necessário. Como colocado por Leclerc, a existência é contingente, mas a maneira como ele apresenta essa contingência não mistura o existente com o que pode ser conhecido:

Um fato é contingente quando nem é necessário nem impossível. A existência das coisas ao nosso redor é contingente. Noutras palavras, numa formulação *de dicto*, qualquer afirmação existencial relativa a um objeto espaço-temporal é contingente. É possível desconhecer se algo existe ou não. Podemos buscar coisas que não existem, como o Eldorado ou a Fonte da Juventude. Por isso, frases como “Ponce de León procurava a Fonte da Juventude na Flórida” não são extensionais (pois não autorizam a generalização existencial) (LECLERC, 2006, p. 117).

²⁸ É bastante intuitivo e aparentemente óbvio pensar que há coisas inacessíveis à cognição humana. Isso não equivale à adoção de uma postura “misteriosista” em relação a uma coisa específica. Dizer que há coisas que não podemos conhecer não as define, nem estabelece que já sabemos de antemão que jamais poderemos conhecer a coisa x ou y. Se não podemos conhecer, então como poderíamos falar sobre elas? A posição misteriosista, por outro lado, estabelece precisamente o que não pode ser conhecido, por exemplo, a mente, a ação correta, a origem da vida etc.

Essa abordagem da existência é bastante apropriada frente a uma perspectiva interpretivista, pois são as coisas circundantes que estabelecem, juntamente com os interlocutores, uma âncora para a linguagem, para a cognição e para as mentes. Nesse sentido, se há um direcionamento das nossas crenças ao acordo coletivo e a uma relação em grande medida confiável com o mundo circundante, isso é formado a partir dos contextos de comunicação específicos. Assim, ao afastar-se de uma busca pelas verdades universais e aproximar-se dos contextos contingentes nos quais as relações de comunicação ocorrem de fato, torna-se possível o estabelecimento de novos pontos de partida para a compreensão da relação entre sujeito e mundo objetivo.

1.3.2. Fechamento causal do mundo físico e estados mentais causais

O problema da causação mental é intrigante. Ele surge porque observamos quotidianamente os nossos estados mentais terem efeitos sobre a realidade física. Porém, a realidade física é nomologicamente fechada. Então, como poderiam nossos estados mentais terem influência sobre ela? Por exemplo, em uma formulação cartesiana, considerando que Descartes apresenta mente e corpo como substâncias distintas, a questão seria compreender como duas substâncias distintas poderiam interagir.

Porém, mesmo depois que se tornou uma espécie de consenso entre os filósofos de que não há uma substância não-física, de modo que resta uma esperança de que o mental possa ser estudado com os métodos das ciências naturais, o problema não foi superado. Como colocado por Amaral (2001), “a pergunta crucial parece ser a seguinte: dado que para cada estado mental há um estado físico correspondente que o ocasiona, qual o papel causal do primeiro face à influência causal óbvia do segundo na ocorrência de outros estados físicos (ou mentais)?”

A questão da causação mental suscita tantos debates que alguns filósofos a consideram como o “coração do problema mente-corpo” (ROBB; HEIL, 2013, p.3). A dificuldade acerca de como conciliar o fechamento causal do mundo físico e as nossas intuições acerca do papel causal de nossas crenças, desejos e decisões deve ser enfrentada. Seguindo Kant, Davidson argumenta que essa questão é importante para que se possa conciliar a autonomia da ação humana com a necessidade do mundo natural (DAVIDSON, 1970/1994b, p. 5), e não se deve desviar dela.

O materialismo não reduutivo que acompanha a perspectiva interpretivista permite uma conciliação entre o fechamento causal do mundo físico e a ideia de que os eventos mentais (ao

menos alguns deles) fazem parte da cadeia causal natural. Os eventos mentais não são algo ontologicamente além dos eventos físicos, mas só podem ser descritos por meio de um vocabulário próprio de nível alto.

Por um lado, a influência que nossos estados mentais exercem sobre o mundo é fundamental para que nos vejamos como agentes. Nossos pensamentos e sensações são importantes explicações para nossos comportamentos. Por exemplo: ao fazer uma comida, a experimento e a avaliação que faço do sabor que sinto explica o fato de eu adicionar mais sal à panela. Da mesma forma, minha vontade de ouvir uma música faz com que eu ligue a caixa de som. Os eventos mentais, portanto, parecem produzir uma série de alterações comportamentais com efeitos complexos sobre o mundo exterior. A relação entre nossos estados mentais e o mundo exterior, aparentemente, é necessária para que possamos construir uma imagem de nós mesmos e das outras pessoas.

Por outro lado, há a crença amplamente difundida e aceita de que o mundo natural é nomologicamente fechado. Nesse sentido, como poderiam estados mentais possuir poderes causais comportamentais se não fossem parte do mundo físico? Ora, se os estados mentais possuem efeitos sobre o mundo físico, então eles devem compartilhar de uma mesma ontologia com a física. Todos os eventos físicos parecem ter uma causa física suficiente. Negar isso seria admitir que há algo sobrenatural capaz de provocar alterações no mundo natural e seria difícil aceitar essa possibilidade pois, novamente, ficamos diante do problema: como uma substância sobrenatural poderia provocar alterações no mundo natural?

Alguns autores tentariam acomodar sua visão de mundo apelando para o epifenomenalismo. Porém, essa é uma posição anti-intuitiva demais. As nossas experiências não poderiam ter qualquer influência sobre nosso comportamento. Geralmente, o epifenomenalismo acaba sendo a saída para quem defende os *qualia* como estados mentais não relacionais, mas parece um preço alto demais a se pagar. Conforme McLaughlin:

O epifenomenalismo é uma doutrina verdadeiramente assombrosa. Se ela é verdadeira, então nenhuma dor poderia jamais ser a causa da retração de nossos músculos, nem seria possível que alguma coisa parecendo vermelha para nós fosse jamais a causa do nosso pensamento de que ela é vermelha. Uma dor de cabeça persistente jamais poderia ser a causa de um mau humor (McLAUGHLIN, 1995, p. 277, tradução nossa).

Yudkowsky (2015, p. 1023, tradução nossa) não consegue compreender a razão pela qual alguns filósofos adotam uma postura epifenomenalista:

Quanto a mim, eu diria que se você postula uma propriedade da consciência misteriosa, separada, adicional, inerentemente mental acima e além das posições e velocidades, então, neste ponto, você já ofereceu o seu pescoço tanto quanto

possível. Postular essa coisa de consciência e depois postular que *ela não faz nada* - pelo amor de gatinhos fofos, *por quê?* (YUDKOWSKY, 2015, p. 1023, tradução nossa)

A posição interpretivista de Davidson e de Dennett está mais próxima do senso comum nesse sentido. Eles negam o epifenomenalismo e sustentam que as ações humanas são causadas por eventos mentais e estes estão diretamente relacionados aos eventos físicos. Davidson, por exemplo, entende que os eventos físicos causam crenças e estas, por sua vez, causam comportamentos²⁹:

(...) [S]e alguém afundou o Bismarck, então vários eventos mentais tais como percepções, observações, cálculos, juízos, decisões, ações intencionais e mudanças de crenças desempenharam um papel causal no afundamento de Bismarck. Em particular, eu insisto que o fato de alguém ter afundado o Bismarck implica em que tal pessoa moveu seu corpo de maneira que foi causada por eventos mentais de certa sorte e que esse movimento corporal foi, por sua vez, a causa do afundamento do Bismarck. A percepção ilustra como a causalidade pode correr do físico para o mental: se um homem percebe que um navio está se aproximando, então a aproximação do navio deve ter causado nele a crença de que o navio está se aproximando. (DAVIDSON, 1970/1994b, p.7, tradução nossa).

Dennett apresenta um argumento interessante contra o epifenomenalismo. O argumento é que, no mundo da causalidade material, tudo permanece igual independentemente da existência de qualquer coisa que seja epifenomênica. Por isso, a crença em epifenômenos não tem fundamento, pois ela não pode ser causada pelo próprio epifenômeno. Por exemplo, a crença em que há *qualia* epifenomênicos seria causalmente desconectada dos próprios *qualia*.

Se os *qualia* são epifenomênicos no sentido filosófico padrão, sua ocorrência não pode explicar a maneira como as coisas acontecem (no mundo material) uma vez que, por definição, as coisas aconteceriam exatamente da mesma maneira sem eles. Não poderia haver uma razão empírica, então, para acreditar em epifenômeno. Poderia haver outra sorte de razão para afirmar a existência deles? Que sorte de razão? Uma razão *a priori*, presumivelmente. Mas qual? Ninguém jamais ofereceu uma - boa, má ou indiferente - que eu tenha visto. Se alguém quiser objetar que eu estou sendo um “verificacionista” sobre esses epifenômenos, eu respondo: Todo mundo não é verificacionista sobre essa sorte de afirmação? Considere, por exemplo, a hipótese de que haja quatorze “gremlins” epifenomênicos em cada cilindro de um motor de combustão interna. Esses “gremlins” não têm massa, não têm energia, não têm propriedades físicas; eles não fazem o motor correr mais suave ou mais áspero, mais rápido ou mais devagar. Não há e não poderia haver evidência empírica de sua presença, e nenhuma maneira empírica em princípio de distinguir essa hipótese de suas rivais: há doze, ou treze, ou quinze... “gremlins”. Mas qual é princípio pelo qual se defende que tal absurdo indiscriminado seja recusado? Um princípio

²⁹ Aqui não é possível imiscuir-se de mencionar a crítica de Kim (1993) a Davidson. De acordo com Kim, Davidson não é capaz de evitar o epifenomenalismo, pois os estados mentais estariam submetidos à causalidade material apenas enquanto estados físicos. Porém, essa crítica ainda parte de uma ruptura entre os aspectos físicos e mentais da natureza. Para Davidson, estados mentais são estados físicos descritos de outra forma. Ver: seções 1.4; 4.3.3 e 4.5.

verificacionista, ou apenas simples senso comum? (DENNETT, 1991b, p. 403-404, (2015, p. 1023, tradução nossa)

O argumento de Dennett contra o epifenomenalismo evidencia o caráter infalsificável da teoria. Assim, a ideia de estados mentais não relacionais é rejeitada por ambos os autores: Davidson e Dennett. Para eles, as ações dos sujeitos podem ser explicadas a partir da atribuição de estados mentais aos sujeitos, ou seja, o que explica os comportamentos são os estados mentais. Isso não significa que cada comportamento de um agente tenha uma explicação específica em termos de estados mentais. Há diferentes eventos mentais que poderiam explicar uma mesma ação. Há diversas razões pelas quais eu posso me levantar para acender a luz. Por exemplo, eu posso desejar ler e crer que seria melhor fazer isso com a luz acesa. Além disso, posso crer que ler com a luz apagada faça mal à minha saúde ocular e também posso desejar que uma amiga, ao passar pela rua, veja minha luz acesa, saiba que eu estou em casa e venha me fazer uma visita. Um mesmo evento, portanto, poderia ser descrito das mais diversas maneiras e não há como selecionar de maneira regular e objetiva uma ou algumas atitudes proposicionais para explicar a mesma ação. Essa é a noção de indeterminação da interpretação, que será melhor desenvolvida no último capítulo. Davidson e Dennett compreendem a indeterminação de formas diferentes, mas ambos aceitam a tese herdada de Quine.

1.3.3. Caráter anômalo do mental

Diante da indeterminação da interpretação, Davidson explica que, embora os motivos do agente sejam as causas das ações, não há leis estritas que expliquem a regularidade causal na relação entre os eventos mentais e as ações humanas. Nesse sentido, ainda que o mental faça parte da cadeia causal da natureza, afirma Davidson, ele é anômalo. Não é possível nem relacionar o mental e o físico por meio de leis estritas (assim como ele acredita que seriam idealmente as leis da física), nem explicar o próprio mental por meio de leis estritas, ainda que tais leis fossem formuladas em um vocabulário puramente mental. Compreender um evento mental é compreender o seu papel em meio a uma rede de outros eventos mentais, e nisso se constitui o holismo do mental, como parte de sua anomalia. Ao falar sobre a impossibilidade de redução do mental tanto em um vocabulário comportamental quanto por meio de leis estritas que o vinculem ao físico, Davidson aponta para o fracasso do behaviorismo e argumenta:

(...) [S]e estivéssemos prestes a encontrar uma proposição aberta expressa em termos comportamentais e exatamente coextensiva com algum predicado mental, nada nos poderia persuadir razoavelmente que a tivéssemos encontrado. Sabemos demais sobre o pensamento e o comportamento para dar crédito a enunciados exatos e universais que os vinculem. As crenças e os desejos aparecem no comportamento apenas quando modificados por posteriores crenças e desejos, atitudes e expectativas, sem limite. Está claro que esse holismo do âmbito mental é um indicio tanto da autonomia como do caráter anômalo do mental. (DAVIDSON, 1970/1994b, p. 35, tradução nossa).

Um comportamento ou um estado mental, portanto, só podem ser explicados a partir de suas relações com outros comportamentos e outros estados mentais. Para podermos explicar e prever o comportamento de um indivíduo, é necessário, antes de qualquer coisa, que lhe atribuamos racionalidade e coerência. “É sempre possível, é claro, melhorar a compreensão que uma pessoa tem de outrem” (DAVIDSON, 1995, p. 232, tradução nossa). Para isso, precisamos de mais dados acerca do sujeito ou de apelarmos racionalmente para nossa imaginação ao conferir coerência às ações do sujeito. Ainda assim, nossas explicações e previsões serão sempre imprecisas, de modo a nunca poderem se reduzir em termos de leis estritas.

O monismo anômalo de Davidson decorre dos três princípios³⁰: O princípio da interação causal, segundo o qual há eventos mentais que têm causas físicas; o princípio do caráter nomológico da causalidade, ou seja, que eventos que mantêm relação de causa e efeito são regidos por leis estritas e o anomalismo do mental, que eventos mentais e físicos não se relacionam por meio de leis estritas.

Cada vez que se descreve um evento mental, é possível fazer uma descrição física desse mesmo evento. Isso vale para todos os eventos mentais, mas não para todos os eventos físicos, de modo que muitos eventos físicos não possuem uma descrição mental. Assim, ao mesmo tempo em que assegura o fechamento causal do mundo físico, sem postular entidades ou propriedades mentais distintas, Davidson preserva as referências da psicologia de senso comum a intenções, desejos e crenças, por exemplo.

O monismo anômalo é uma posição segundo a qual a psicologia não pode ser epistemologicamente reduzida à física, embora possua a mesma ontologia da física. Há a influência do físico sobre o mental, por exemplo, quando o movimento das folhas da palmeira me leva à crença de que está ventando. Há também a influência do mental sobre o físico, pois a crença em que venta me leva a abrir a janela e deixar o vento entrar em minha casa.

Os princípios (1) da interação causal e (2) do caráter nomológico da causalidade parecem contraditórios com o princípio (3) do anomalismo do mental. O primeiro enuncia que

³⁰ Ver: Introdução.

há interação causal entre o físico e o mental. O segundo, que onde houver causalidade há leis estritas. O último, que não há leis estritas relacionando o físico e o mental. Davidson está empenhado em conciliar esses três princípios porque, dessa forma, é possível preservar o fechamento causal do mundo físico e, ao mesmo tempo, garantir a autonomia da ação humana.

Tomando, então, como pressupostos esses três princípios, somos levados a uma teoria da identidade de eventos particulares (“token-identity”). Suponha que um determinado evento mental *M* seja causa de um evento físico *P* (o princípio da interação causal permite isso). Nesse caso, pelo princípio do caráter nomológico da causalidade, esta interação é um exemplo de uma lei estrita. Ora, essa lei não pode ser mental nem psicofísica, pois o princípio do anomalismo do mental proibiria. Assim, ela deve ser uma lei física. Logo, o evento mental *M* é também um evento físico. Em outras palavras, o evento *M*, se tem relações causais com o evento *P*, deve possuir propriedades físicas e, apenas pela via dessas propriedades, é que se pode correlacionar nomologicamente *M* e *P*. Embora *M* seja um evento mental, como consequência dos princípios (1), (2) e (3), ele deve ter uma descrição física a partir da qual se relacione com *P*. Temos, portanto, uma identidade entre um evento mental particular e um evento físico particular. A mesma linha de argumentação seria seguida para um evento físico que fosse causa de um evento mental.

Nesse sentido, entender Davidson requer uma distinção entre ontologia e epistemologia. Sua ontologia é monista e os eventos mentais são, de fato, eventos físicos. Porém, uma explicação adequada de um evento mental recorre às razões desse evento. Não conseguiremos explicar apropriadamente as causas da ação de alguém se utilizarmos um vocabulário da física. A explicação adequada deve apelar para um vocabulário que recorra a outros eventos mentais. Por exemplo, a causa de João pegar seu guarda-chuva antes de sair de casa foi sua crença em que iria chover que, por sua vez, foi causada pela percepção de muitas nuvens cinzentas e a crença adicional em que tais nuvens precederiam a chuva. A causação, por sua vez, independe da maneira como o evento seja descrito e se submete a leis estritas. O fato de não podermos explicar um evento mental por meio de sua participação na cadeia causal do mundo natural não significa que ele não participe dessa cadeia.

Do ponto de vista epistemológico, porém, os eventos mentais como tipo não podem ser adequadamente explicados por meio de leis estritas da física. As explicações relevantes acerca dos eventos mentais incluem outros eventos mentais. A identidade entre o mental e o físico, portanto, ocorre entre eventos mentais e físicos particulares. Não é possível fazer

generalizações estritas que vinculem tipos de eventos mentais a tipos de eventos físicos. Por exemplo, não é possível identificar como tipo a crença x com a sinapse cerebral y .

1.4. Um bom reducionismo

Foi dito que a resposta de Davidson ao problema da causalção mental é que, para cada evento ao qual se pode dar uma explicação em termos mentais, pode-se também dar uma explicação em termos físicos, pois eventos mentais são eventos físicos. A causalidade estrita funcionaria perfeitamente se tivéssemos condições epistêmicas para observar todos os aspectos físicos relevantes na causalidade, ela estaria clara diante dos nossos olhos. Porém, isso não poderá nos dizer as razões dos comportamentos das pessoas. Para isso, precisamos de uma descrição daquele mesmo evento em termos intencionais.

Acontece que tentar explicar o comportamento de uma pessoa em termos físicos exigiria o controle de uma série de fatores, dentro e fora do corpo daquela pessoa. Como criaturas epistemologicamente limitadas, não temos condições de fazer isso. Todo o período da vida de um cientista cognitivo talvez não seja suficiente para explicar a única ação de um agente de atravessar uma rua, pois isso envolve inúmeros fatores que precisariam ser controlados. Como, então, se poderia explicar, por exemplo, uma aula de filosofia a partir de uma descrição física?

Não é que a descrição física não exista. Ela existe, a aula de filosofia não é um evento sobrenatural, ela ocorre no mundo. Porém, essa descrição física não é capaz de esclarecer o que está acontecendo ali. Para esclarecer a aula de filosofia, é preciso atribuir ao professor e aos alunos uma série de estados intencionais. Esse é espírito do materialismo não reducionista de Davidson (mas que também está presente em Dennett): do ponto de vista ontológico, um monismo. Do ponto de vista epistemológico, diferentes níveis de explicação.

Acontece que a posição de Davidson às vezes não é considerada satisfatória, pois segundo Amaral (2001), Davidson “confere aos eventos mentais poder causal em um mundo físico fechado causalmente, sem no entanto salvar um papel causal atuante para as propriedades, ou tipos, mentais, dos quais os eventos mentais são ocorrências”. Essa é a crítica de Kim ao monismo anômalo: de que os eventos mentais não entrariam na corrente causal da qual os eventos físicos fazem parte.

A crítica de Kim, contudo, parece partir de um princípio: o de que eventos mentais e eventos físicos são, de alguma forma, distintos, nem que essa distinção esteja no nível de propriedades. Em outras palavras, parece haver um dualismo subjacente a essa crítica. Na

concepção de Davidson, eventos mentais são eventos físicos sob uma descrição intencional, justamente por isso, eles não quebram o fechamento nomológico do mundo físico. Se fossem algo diferente exercendo causalidade sobre o mundo físico, o problema cartesiano da interação seria novamente colocado.

O fechamento causal do mundo físico forma a camada básica da nossa ontologia. No mais, há diversas formas de explicar o mundo que nos cerca. É provável que muitas das questões em filosofia da mente tenham uma só origem: um reducionismo exagerado que busca na física (como disciplina) uma explicação para todas as coisas que existem. Do fato de se defender um monismo ontológico não se deriva que a disciplina da física terá recursos para explicar todas as coisas. É isso que Davidson defende em “Thinking Causes” (1993/2005b): o monismo ontológico, segundo ele, não implica em reducionismo epistemológico.

Por exemplo, a química como disciplina conta com termos, instrumentos e métodos que não fazem parte da física como disciplina. Nem por isso se diria que os eventos químicos estão fora da cadeia causal da física ou que é preciso explicar a causação química. Da mesma forma, a biologia precisa de seus próprios recursos para compreender os seres vivos, mas isso não significa que os seres vivos sejam algo não físico. Assim, se a física não conseguir explicar o comportamento, a linguagem e as sociedades, está tudo dentro do esperado. Acontece que a física é uma disciplina com impressionantes poderes explicativos e preditivos, mas por mais complexos que sejam seus objetos de pesquisa, eles não têm o grau de complexidade das relações humanas.

O reducionismo epistemológico exagerado que tenta explicar tudo com um vocabulário da física provavelmente acaba por gerar uma grande frustração. Ao perceberem as dificuldades dessa empreitada, os teóricos do mental caem ou em uma posição epifenomenalista ou em uma posição eliminativista³¹, segundo a qual o vocabulário deve ser eliminado e substituído por um vocabulário da física. Ora, se isso for feito, como poderemos explicar e prever os comportamentos de nossos amigos e inimigos?

Esse tipo de reducionismo excessivo que busca explicar tudo a partir da física é denominado por Dennett de “ganancioso” [“greedy”]. Para Dennett, nada há de errado em tentar explicar as coisas a partir de suas partes mais simples, nem em buscar explicações de alcance amplo (como ele pensa que é o algoritmo em funcionamento no processo evolutivo). Essa postura seria, para ele, totalmente defensável³². O problema que acaba por levar alguns

³¹ O eliminativismo afirma que a nossa compreensão de senso comum dos estados mentais está completamente equivocada, que os estados mentais dos quais falamos quotidianamente não existem. Nesse sentido propõe que o vocabulário da psicologia de senso comum seja eliminado do discurso científico.

³² “Reduacionismo” já é um termo vilão, usado pelos filósofos geralmente quando querem atacar alguém: “-

autores a desistir do reducionismo e apelar para uma postura totalmente desconectada de qualquer base material é o reducionismo apressado demais.

[...] na sua ansiedade de barganhar, no seu zelo em explicar muita coisa depressa demais, cientistas e filósofos frequentemente subestimam as complexidades, tentando pular camadas ou níveis inteiros de teorias, na pressa de fixar tudo muito bem e com segurança às bases. Esse é o pecado do reducionismo ganancioso, mas observe que só quando o zelo excessivo leva à falsificação do fenômeno que devemos condená-lo (DENNETT, 1995/1998a, p. 86).

Dois exemplos de frustração perante o reducionismo ganancioso são, por um lado, o behaviorismo radical, quando propõe a substituição do vocabulário da psicologia de senso comum por um vocabulário completamente comportamental e, por outro lado, as posturas misteriosistas, segundo a qual a irreduzibilidade do mental implica em uma possibilidade de explicá-lo a partir de qualquer perspectiva que seja objetiva. Na proposta de Dennett, o vocabulário da psicologia de senso comum não deve ser alterado, ele está muito bem assim. Porém, ele se refere a algo que pode ser estudado a partir de uma perspectiva externa: a perspectiva do intérprete.

O problema da redução vocabular é que ela não leva em conta a importância explicativa da psicologia de senso comum. Não podemos abrir mão da psicologia de senso comum para nada que esteja relacionado ao humano. Porém, isso não significa que a mente seja algo que tenha surgido por um acaso inexplicável do destino dentro de alguns sujeitos (os humanos, pelo menos) e que não seja possível explicá-la. Os dois erros provêm de um reducionismo exagerado demais.

O bom reducionismo, segundo Dennett, funciona em forma de cascata. Os eventos de nível superior podem ter uma explicação a partir de uma decomposição em partes mais simples, esses sendo decompostos em partes mais simples e assim sucessivamente, de cima para baixo. De baixo para cima, a explicação seria em forma de uma pirâmide composta de andares. Os andares superiores sempre são apoiados nos andares inferiores. Tudo isso é feito aos poucos, passo a passo. As explicações de nível mais alto podem parecer miraculosas, mas elas têm uma razão para serem assim: é uma forma de lidar com situações que exigem tomadas de decisão rápidas e não tão precisas, como são as situações enfrentadas pela psicologia de senso comum.

Uma metáfora usada por Dennett (1995/1998a, p. 411) é a seguinte: “reducionistas pensam que tudo na natureza pode ser explicado sem *skyhooks*; reducionistas gananciosos

Fulano é reducionista!” Mas Dennett não pensa assim. O reducionismo é bom, desde que seja cuidadoso e comedido.

pensam que tudo pode ser explicado sem gruas”. “Skyhooks” seriam supostos ganchos que viriam do céu trazendo algo, sem terem nada sobre que se apoiarem. Eles se opõem às gruas, que se apoiam no chão para poderem elevar algo até o céu. As explicações que se baseiam em fenômenos anteriores e mais simples, então, não deixam de ser bem-vindas.

Por exemplo, se queremos explicar a intencionalidade, convém partirmos do pressuposto de que ela não surge do nada. O caminho mais apropriado seria contar uma história que permita compreender como a intencionalidade surge a partir do que não é intencional, uma história de base evolutiva. Se isso for possível, evitam-se as grandes lacunas explicativas assim como se torna desnecessário o recurso ao milagre. Dennett oferece uma versão especulativa dessa história, sua “torre de gerar e testar”³³.

Dennett acredita que a teoria da evolução é um recurso muito importante para podermos traçar esse caminho, por duas razões. A primeira, é bastante óbvia: a evolução das espécies construiu a complexidade a partir da simplicidade, gradualmente, por meio de um processo que o autor apresenta como um algoritmo, um algoritmo simples, mas capaz de produzir resultados grandiosos.

Por um lado, poucos filósofos tentaram negar o fato óbvio: os seres humanos são produtos da evolução, e sua capacidade de falar, e de dizer algo com significado (no sentido relevante), portanto, se deve a uma sequência de adaptações específicas não partilhadas com outros produtos da evolução. Por outro lado, os filósofos têm relutado em pensar na hipótese de que o pensamento evolutivo possa esclarecer os seus problemas específicos sobre como é que as palavras, e suas fontes e destinos nas mentes das pessoas ou cérebros, têm significados (DENNETT, 1995/1998a, p. 420).

A outra razão é que, na concepção de Dennett, o algoritmo que “roda” a evolução das espécies também “roda” uma série de outros processos existentes. Algoritmos são processos simples e mecânicos, mas alguns deles podem produzir resultados surpreendentes. Um algoritmo, segundo ele, possui três características: (1) *neutralidade de substrato*, ou seja, o algoritmo não depende de materiais específicos para que ele possa funcionar, precisa apenas de uma estrutura física que permita o seu funcionamento; (2) *irracionalidade subjacente*, quer dizer, o algoritmo é executado por meio de etapas simples e estúpidas e (3) *resultados garantidos*, se as etapas do algoritmo forem corretamente executadas, haverá um resultado.

³³ Ver: seção 3.1.7.

1.4.1. Sobre o algoritmo evolutivo³⁴

Os programas de computadores, muitos problemas matemáticos e as receitas de bolo podem ser algoritmos. Aqui, temos exemplos de algoritmos cujos resultados são previsíveis e controláveis. Alguns processos algorítmicos, contudo, podem ser bem pouco previsíveis. A característica (3) afirma que os resultados são garantidos, o que não significa que sejam resultados predeterminados. Muitas vezes, não temos como saber de antemão que resultados seriam esses.

Conforme o exemplo de Dennett, um torneio qualquer pode ser descrito como um algoritmo. Em muitos torneios há dois elementos envolvidos: habilidade dos jogadores e sorte. Num torneio de tênis, espera-se que o melhor tenista vença, mas alguns fatores incontroláveis podem levá-lo a perder o torneio. Quanto maior a quantidade de torneios de que o jogador participar, maiores as suas chances de mostrar a sua técnica, pois, nesse tipo de jogo, a habilidade tende a se sobrepor à sorte.

Por outro lado, em um torneio de cara ou coroa, apenas o elemento sorte estará presente. Assim, não há melhor jogador, apenas o fato de alguém ter ganhado aquele torneio específico. Quanto mais campeonatos forem feitos com os mesmos jogadores, menores as chances de que alguém se sobressaia aos demais. O algoritmo dos torneios tem um resultado: é um vencedor. Porém, não garante que o vencedor seja o mais habilidoso.

Dennett (1995/1998a, p. 57-8) pede que imaginemos um campeonato no qual os jogadores são os mesmos, mas as condições se modificam. Por exemplo, o primeiro jogo é de xadrez. Os vencedores da primeira etapa jogam tênis na segunda. Na terceira, os vencedores jogam bilhar, depois os vencedores jogam golfe, e assim sucessivamente, trocando os jogos, até que saia o vencedor. Nesse caso, embora o vencedor tenha se saído bem em cada uma das etapas, isso não garante que ele seja o melhor em nenhuma delas. Ele pode ter eliminado na primeira etapa um ótimo jogador de bilhar que teria ganhado dele posteriormente, mas que não teve a oportunidade de chegar ao jogo de bilhar. O vencedor desse torneio, ao final, terá contado com a sorte de ter eliminado previamente aqueles que seriam seus grandes rivais nas

³⁴ A ideia do algoritmo evolutivo ainda permanece em Dennett (2017b), porém, com o desenvolvimento das pesquisas científicas, especialmente com o trabalho de Godfrey-Smith (2007), Dennett percebe que vários outros processos bem mais complexos também entram em cena quando se trata de evolução. Ainda assim, a compreensão desse algoritmo é capaz de esclarecer aspectos importantes do processo evolutivo, em especial, seu caráter contingente e local e sua capacidade de produzir grandes resultados por meio de pequenos passos simples. Esta seção se apoia em (FAGUNDES, 2009).

etapas seguintes. Como a evolução tem um caráter local, ela tem aspectos em comum com esse jogo múltiplo.

Dados esses três elementos, o algoritmo evolutivo começará a produzir resultados, independentemente das especificidades do processo envolvido: variação, retenção (ou hereditariedade) e seleção (DENNETT, 1995/1998a, p. 357; BLACKMORE, 1999, p. 10). Em primeiro lugar, deve haver uma abundância de elementos diferentes entre si (variação). Essas variantes são capazes de produzir cópias fiéis de si mesmas (hereditariedade), embora ocasionalmente ocorram erros de cópia, gerando variação. Além disso, algumas dessas variantes estão mais aptas a sobreviver e se reproduzir do que as outras, isto é, variantes que produzem mais cópias do que outras.

A ideia de Darwin para explicar como ocorre a evolução das espécies é suficientemente abstrata para ter sido elaborada sem que existisse qualquer conhecimento acerca dos mecanismos específicos envolvidos no processo de variação e retenção seletiva. Não havia qualquer conhecimento acerca do gene, a unidade preservadora da hereditariedade. Ainda assim, a teoria foi construída. O algoritmo evolutivo construído a partir da ideia de Darwin, segundo Dennett, entra em funcionamento sempre que as condições abstratas forem preenchidas.

A ideia não é dennettiana. Como colocado por Campbell (1960), esse processo pode ter ocorrido sucessivamente em diversos níveis, não apenas no nível orgânico, como também no nível da aprendizagem e criatividade humanas. Posteriormente, na mesma linha de Campbell, Popper (1972/1992) também o aplica ao desenvolvimento científico, afirmando que o crescimento do conhecimento se dá por um processo de seleção natural de hipóteses. As hipóteses que produzimos são submetidas a um processo seletivo. Permanecem em voga aquelas que estão aptas a resistir a esse processo, pelas mais diversas razões. Aqui, as variantes envolvidas no processo não são os genes, mas sim as hipóteses.

Dennett (1996/1997; 1998) argumenta que esse processo pode ocorrer em níveis sucessivos de maneira que aspectos complexos do mundo, como a aprendizagem, a linguagem e o desenvolvimento da cultura possam ser derivados de uma série de etapas simples ocorrendo continuamente ao longo de grandes intervalos. Assim, ele propõe uma organização dos seres vivos em camadas hierárquicas, umas sobre as outras, formando uma torre, sendo que as camadas superiores possuem nível de organização mais complexo. À medida que a torre cresce, as criaturas se tornam mais aptas para prever e evitar os possíveis desafios que lhes são colocados pelo ambiente. As criaturas no topo da torre são muito complexas, mas resultaram do mesmo processo algorítmico simples e gradual que gerou as criaturas mais

simples. Dennett admite que o processo descrito por ele seja excessivamente simplificado, mas mostra como o algoritmo pode ocorrer em diferentes níveis até o surgimento de animais possuidores de mente, como nós.

Se é possível a aplicação do algoritmo evolutivo a todos esses processos, estamos diante de um mecanismo realmente poderoso e altamente explicativo. Porém, vale esclarecer que isso não é um elemento tão central ao interpretivismo. É só para mostrar a relevância que Dennett dá ao pensamento evolutivo. O mais importante, aqui, é levar em consideração que (1) a intencionalidade não surge do nada dentro do sujeito; (2) o reducionismo ganancioso vai desembocar em problemas, pois não há apenas um nível de explicação da realidade; (3) também não há apenas dois níveis de explicação da realidade (o físico e o mental), mas diversos níveis sobrepostos em camadas; (4) redução ontológica não implica em redução epistemológica.

Quanto a esse último item, o que se quer dizer é que é possível ter uma perspectiva monista da realidade, mas isso não significa que todas as coisas que acontecem ao nosso redor devam ser explicadas por meio de um vocabulário da física, da química ou da biologia. Precisamos dos vocabulários das ciências de níveis mais altos. Precisamos, sobretudo, de vocabulários que ajudem a compreender a mente e só é possível compreendê-la na medida em que se é possível falar sobre ela. Mas por que postular entidades subjetivas e inefáveis? Não vale a pena empenhar-se mais para tentar outro caminho?

1.5. Intencionalidade

Serão necessárias aqui algumas palavras sobre “intencionalidade”, pois é uma noção importante para o interpretivismo. Diferente do uso comum do termo, que significa desejo, vontade, pretensão ou propósito, “intencionalidade” tem uso técnico em filosofia da mente³⁵.

³⁵ Há outro termo técnico que não deve ser confundido: “intensionalidade” (com “s”, termo derivado da palavra latina *intentio*, ou o conceito). Em filosofia da linguagem se distingue intensão e extensão de um termo. Extensão é o conjunto dos objetos capturados por aquele conceito, enquanto a intensão é o modo particular como aqueles objetos são capturados, também entendido como o significado. Tome, por exemplo, a noção de “animais que têm rins” e de “animais que têm coração” – nesse caso, temos a mesma extensão para diferentes intensões. Em muitos casos, dentro de uma proposição, é possível fazer a substituição *salva veritate* de termos que possuem a mesma extensão, são os contextos extensionais, ou translúcidos. Por exemplo, se temos a frase verdadeira “Aristóteles nasceu em Estagira”, podemos substituir “Aristóteles” pelo sinônimo “O mestre de Alexandre” e teremos mantido o valor de verdade na frase “O mestre de Alexandre nasceu em Estagira”. Contudo, há situações em que esse tipo de substituição não é possível, são os contextos intensionais (com “s”), ou contextos opacos. As descrições de estados intensionais (com “c”) são exemplos de contextos intensionais (com “s”). Por exemplo: não podemos fazer a mesma substituição *salva veritate* em “Marcos pensa que Aristóteles nasceu em Estagira”. Os contextos intensionais são aqueles no qual a verdade depende do significado, e não apenas da referência de suas palavras componentes.

Está relacionado ao uso comum, mas tem uma abrangência maior. Em filosofia, significa a qualidade de estar dirigido para algo, possuída por pelo menos alguns estados mentais: “Alguma coisa exibe intencionalidade se sua competência é de algum modo sobre outra coisa” (DENNETT, 1996/1997, p. 39, grifo do autor). Conforme Jacob (2014), deriva do verbo latino *intendere*, que significa justamente estar direcionado para um objeto ou coisa. O debate sobre intencionalidade foi introduzido na contemporaneidade por Brentano, que recuperou o termo da escolástica.

Por exemplo, no sentido comum, quando alguém comete um erro, as pessoas podem discutir a respeito de se a pessoa tinha a intenção de fazer aquilo. Este é um sentido mais restrito do termo. É certo que, aqui, a intenção está direcionada a algo, isto é, àquele erro cometido. Porém, o sentido filosófico é mais amplo. A memória e os estados emocionais exprimíveis em termos de atitudes proposicionais, por exemplo, são intencionais no sentido filosófico, pois estão direcionados a algo, mas não são intencionais no sentido comum. Quanto às atitudes proposicionais, elas promovem justamente esse direcionamento de um sujeito a uma proposição.

Algo que se poderia discutir é se todos os estados mentais possuem intencionalidade, pois alguns autores consideram que há estados mentais que são intencionais e outros que não. Para eles, estados mentais que poderiam estar disponíveis para participar do controle comportamental seriam considerados intencionais, mas haveria estados mentais não disposicionais, como os *qualia*. Ned Block (1995), por exemplo, faz uma diferenciação entre o que é fenomênico e o que é disposicional na mente humana. Chalmers (1996), com seu dualismo de propriedades, também segue por esse caminho.

Dennett (1995/1998a, por exemplo), por outro lado, recusa essa diferença, considerando todos os estados mentais como intencionais. Leclerc (2017) também considera todos os estados mentais intencionais e cada estado mental, para ele, é disposicional. Ele chama a atenção para dois tipos de intencionalidade: vertical e horizontal. A vertical consiste em uma flecha apontada para um objeto, essa é uma metáfora comum para o direcionamento que caracteriza a intencionalidade. A horizontal consiste na continuidade temporal da experiência, seu tecido e consistência. Para ele, a continuidade do agente consiste de uma série de disposições que permanecem. Assim, a intencionalidade também teria a característica da disposicionalidade. Crane (1998) defende que todos os estados mentais e apenas eles são intencionais, o que torna a intencionalidade “a marca do mental”, como defenderia Brentano³⁶.

³⁶ Essa é a terceira tese de Brentano. A primeira é que é constitutivo dos estados mentais que eles são direcionados a coisas diferentes deles mesmos. A segunda é que as coisas às quais os estados mentais se

Porém, para Crane, alguns estados mentais possuiriam propriedades não intencionais³⁷.

No caso de Davidson (1984/2001g), ele parece evitar um pouco essa questão da diferenciação entre estados mentais, dizendo simplesmente que não tratará das dores e outras sensações, consideradas por muitos como aspecto central do mental e que estará preocupado apenas com as atitudes proposicionais – ainda que isso seja uma abordagem parcial. Porém, posteriormente ele parece assumir explicitamente o caráter disposicional das sensações. “Perceber que está nevando, sob circunstâncias apropriadas, consiste em que os próprios sentidos causem (do modo apropriado) a crença de que está nevando quando a neve está caindo” (DAVIDSON, 2001, p. xvi, tradução nossa). Portanto, Davidson acaba por recusar a distinção entre estados mentais disposicionais e não disposicionais.

O termo “intencionalidade” tradicionalmente se referia apenas a estados mentais, e a capacidade de se dirigir a objetos externos era atribuída unicamente aos seres humanos. Foi usado, inclusive, como uma forma de diferenciar entidades puramente físicas e as que possuiriam mente. De acordo com a interpretação de Chisholm (1957 *apud* FRÉCHETTE, 2016), uma das teses de Brentano é a da irredutibilidade da intencionalidade, chamada também de “problema de Brentano”: não se pode sair do círculo da intencionalidade. Em outras palavras, não se pode usar termos não intencionais para definir um termo intencional. Conforme Dennett (1995/1998a, p. 341), Quine chega à mesma conclusão quando conclui que não há nenhuma forma de traduzir a linguagem intencional em uma linguagem das ciências físicas. Porém, ao contrário de Brentano, Quine se alia a um behaviorismo radical, concluindo que uma ciência da intencionalidade teria conteúdo vazio. Segundo Fréchette, a tese de Brentano é compreendida por Quine (1960/2010) como parte da tese acerca da indeterminação da tradução, o que levaria, para Quine, a uma impossibilidade de se estudar o vocabulário intencional sob uma perspectiva científica³⁸. Essa tese influencia Davidson, para quem o vocabulário intencional não pode ser reduzido a um vocabulário baseado na física.

Dennett (1995/1998a) é um autor que vai de encontro à concepção tradicional, acreditando que a intencionalidade se estende desde a evolução natural até os computadores, passando, é claro, pelos seres humanos³⁹. Isso não significa que o autor defenda que qualquer

direcionam têm a propriedade da inexistência intencional (CRANE, 1998).

³⁷ O problema é apenas deslocado quando se adota a posição de que estados mentais disposicionais possuem propriedades não disposicionais. No lugar de dividir os estados mentais entre os que são disposicionais e os que não são, dividem-se as propriedades dos estados mentais entre as que são e as que não disposicionais. As dificuldades relativas aos supostos estados mentais não disposicionais são deslocadas para as propriedades mentais não disposicionais. Aparentemente, os mesmos paradoxos que a questão dos *qualia* geram permaneceriam após essa manobra de Crane.

³⁸ Ver seção 1.6.1.

³⁹ Ver seção 3.1.6.

coisa tenha intencionalidade. “Um mapa, um cartaz, um sonho ou canção podem ser sobre Paris, mas Paris não é sobre nada” (DENNETT, 1995/1998a, p. 212). Tomasello (2014) também é um autor que amplia a aplicação da intencionalidade a outros seres, mas defende que a intencionalidade surge na linhagem homínida a partir de um determinado momento na história evolutiva, antes do surgimento do ser humano e do ancestral comum que temos com os grandes símios, pois estes já exibiam intencionalidade. O interessante desse tipo de abordagem é apresentar a intencionalidade como algo que surge gradualmente, a partir do não intencional.

Aparentemente, há muitas coisas que exibem intencionalidade e não são estados intencionais. Um exemplo em que isso se torna bem nítido é o do livro. Ele é sobre alguma coisa, mas não é um estado mental. Alguns autores que entendem a intencionalidade como a marca do mental explicam essa situação fazendo uma diferenciação entre intencionalidade original e derivada. Searle (2002) é um autor que insiste nessa distinção. Para ele, os estados mentais de organismos naturais possuiriam intencionalidade original. O livro, por outro lado, possuiria apenas intencionalidade na medida em que lhe atribuímos essa característica. Essa seria a intencionalidade derivada.

Dennett não aceita essa diferença entre intencionalidade original e derivada. Para ele, a intencionalidade é algo que se caracteriza a partir da interpretação do comportamento de algum ser. Isso acaba trazendo um aspecto menos problemático ao pensamento de Dennett acerca do mental, evitando ao menos dois problemas: o de estabelecer uma linha demarcatória entre intencionalidade original e derivada e o de evitar a concepção segundo a qual a intencionalidade original seria uma espécie de dádiva milagrosa conferida apenas aos seres humanos ou aos animais superiores⁴⁰.

1.5.1. Interpretivismo quotidiano e idealizado

Numa versão do interpretivismo que podemos chamar de quotidiana, nós atribuímos de fato estados intencionais aos outros (sejam pessoas, animais, máquinas etc.) para podermos dar sentido, explicar e prever o seu comportamento. Dessa forma, há uma rede de estados intencionais conectados entre si aos quais recorreremos para podermos interpretar o comportamento dos seres que possuem intencionalidade. Essa atribuição é extremamente útil, a tal ponto que, sem ela, não poderíamos viver. O sucesso que usualmente se obtém nessa

⁴⁰ Ver seção 3.1.4.

atribuição de intencionalidade é um termômetro capaz de nos permitir a conclusão de que os seres aos quais os estados mentais são atribuídos, de modo geral, possuem aproximadamente aqueles estados mentais que nós lhes atribuímos.

Uma maneira mais técnica de caracterizar a posição interpretivista é apresentada por Kriegel (2011). Dentro dessa perspectiva, alguém (ou algo) teria um estado intencional apenas no caso em que um intérprete ideal interpretaria esse alguém (ou algo) como tendo aquele estado intencional. Um intérprete ideal seria alguém completamente informado acerca dos fatos não intencionais, além de ser completamente coerente e jamais fazer inferências dedutivas inválidas.

É verdade que tanto Dennett quanto Davidson recorrem ao intérprete ideal. Dennett é sutil a esse respeito, pois ele está mais focado nos poderes preditivos reais da psicologia do senso comum. Mesmo assim, ele considera como tendo uma crença alguém a quem essa crença seria atribuída sob a melhor interpretação possível. A melhor interpretação possível é, para ele, a que realiza a melhor previsão comportamental. Ora, aqui entra o intérprete ideal, uma vez que nossas interpretações quotidianas são falhas e imperfeitas, apesar de sua taxa bastante alta de sucesso, pois sequer conseguiríamos entender as palavras de nossos colegas se não estivéssemos fazendo excelentes interpretações de seus comportamentos.

A proposta interpretivista é que se compreenda a mente a partir da atribuição de intencionalidade. Aqui, a mente não é mais vista pelo ponto de vista do sujeito, mas sim sob a perspectiva do intérprete. Podemos recorrer ao intérprete ideal para trazer uma caracterização mais técnica, mas também podemos recorrer às nossas experiências quotidianas de interpretação, que talvez sejam mais esclarecedoras e evite os embaraços filosóficos das situações imaginárias e irreais.

A filosofia pode olhar para as relações reais entre as pessoas e com o mundo material e tirar conclusões daí. Isso poderia ser considerado muito ingênuo por alguns filósofos, mas é um material palpável que temos disponível para nossas pesquisas. É preciso tentar e ver as consequências. No mais, o interpretivismo é uma abordagem que traz consigo, além de uma compreensão do mental, também uma compreensão da linguagem. Ainda evita muitos problemas derivados das concepções do mental que partem do pressuposto de uma ruptura entre subjetividade e objetividade.

1.6. O ancestral do interpretivismo

Ao olhar para nossos pensamentos, é comum encontrarmos um discurso interno repleto de coisas às vezes estranhas, às vezes infundadas, ou vergonhosas. Diante disso, a tendência é conceber o mental como algo privado que se inicia dentro do sujeito e se projeta para fora, mas o interpretivismo inverte esse processo. A ordem dos acontecimentos, então, passa a ser: primeiro surge a necessidade de atribuição de estados mentais aos outros sujeitos. Em seguida, essa necessidade é internalizada.

Essa ideia do discurso mental como algo que ocorre de fora para dentro já havia sido desenvolvida por Sellars desde um importante texto de 1956. Nessa mesma época, Wittgenstein (1953/1979) e Ryle (1949/2009) também trazem novos ares para a filosofia da mente e da linguagem. Ryle critica o dualismo cartesiano e propõe uma leitura da mente pelo ponto de vista comportamental. Wittgenstein contribui com o aspecto social da linguagem e com o argumento contra a linguagem privada⁴¹. Esses três autores passam a exercer grande influência no que estaria por vir dentro da tradição analítica. Sellars pretende, seguindo Ryle, oferecer uma concepção do mental compatível com a psicologia científica. Porém, ao contrário de Ryle, ele não quer negar os estados subjetivos, apenas explicá-los. Para ele, os pensamentos são atividades linguísticas *internas*, mas que têm um papel explicativo frente aos comportamentos públicos. Conforme Beisecker (2015), Sellars deve ser considerado muito próximo, se não um interpretivista.

É certo que é necessário, no mínimo, reconhecer que ele deixou o terreno preparado e lançou as sementes para Davidson e Dennett poderem continuar o trabalho. A nossa concepção de mental, para Sellars, pode ser proveniente não de uma introspecção ou de um acesso direto aos estados internos, mas sim de uma estrutura cultural que é anterior à aprendizagem da linguagem pelos indivíduos (RAMSEY, 2016)⁴².

De acordo com Dennett (1987), Sellars possui uma influência ubíqua na filosofia da mente, mas não é citado na mesma medida. A razão porque ele não é tão citado seriam as dificuldades para compreendê-lo que levam os autores a não perceberem que estão falando da

⁴¹ O argumento da linguagem privada mostra que o vocabulário acerca dos estados mentais não é privado no sentido de se referir a algo que só pode ser conhecido pelo falante, isto é, a experiências privadas. Mostra também que a própria ideia de uma linguagem privada é incoerente. (GLOCK, 1996)

⁴² De acordo com Ramsey, Sellars deixa o caminho aberto para o eliminativismo. Mas o eliminativismo costuma propor a substituição do vocabulário da psicologia de senso comum por um vocabulário fisicalista. O interpretivismo é uma posição que está longe de propor essa substituição. Ao contrário, considera o vocabulário da psicologia de senso comum fundamental até mesmo para que a linguagem possa surgir. Nesse sentido, o pensamento de Sellars fundamenta mais do que um caminho na história da filosofia.

mesma coisa que Sellars. Essa ausência de referências diretas a Sellars, Dennett atribui inclusive a si mesmo. O que é importante, segundo ele, é um ponto de acordo fundamental entre Quine e Sellars: o vocabulário intencional é praticamente indispensável como um elemento de interpretação.

Sellars argumenta que o conhecimento empírico não se apoia em fatos isolados. Ao contrário, só é possível a partir do momento em que a pessoa tem conhecimento de uma série de outras coisas. Por exemplo, para dizer que há um livro amarelo a seu lado, a pessoa precisa antes saber que seus sentidos trazem indícios confiáveis acerca do que se passa em seu entorno. Para saber isso, contudo, é preciso ter tido uma série de observações capazes de apoiar essa hipótese. Não há, na concepção de Sellars, conhecimento sobre fatos isolados que se apoie sobre si mesmo. “[...] Ao caracterizar um episódio ou estado como sendo *conhecimento*, nós não estamos dando uma descrição empírica desse episódio ou estado; nós o estamos situando no espaço lógico das razões, de justificar e ser capaz de justificar o que alguém diz” (SELLARS, 1956/1997, p. 76, tradução nossa).

O *mito de Jones* foi proposto como uma alternativa ao que ele chama de mito do dado [“The Myth of The Given”]. O mito do dado consiste na suposição de que há algo anterior a qualquer conceito e a que o sujeito tem acesso direto e privilegiado. Seria um fundamento pré-linguístico para todo conhecimento empírico, como se o mental tivesse um fluxo próprio de informações independente da linguagem, mas sobre o qual estaria erguido todo conhecimento linguístico possível. Sellars pretende mostrar que isso não é necessário, pois é possível compreender a construção do pensamento sem que haja um fundamento último.

Conforme Rorty (1997), Sellars pretende mostrar que é possível aceitar a dúvida wittgensteiniana sobre uma suposta linguagem privada sem precisar aceitar a dúvida ryliana em relação à existência de entidades mentais como pensamentos e impressões sensoriais. Com isso, ele oferece uma abordagem não redutiva do mental em que a mente se insere no mundo juntamente com a linguagem, em um processo gradual. É essa perspectiva que permite a compreensão de Sellars como interpretivista. Para ele, os acontecimentos internos combinam a privacidade e a intersubjetividade, características do mental.

Inicia-se o mito de Jones com um pedido ao leitor para supor uma comunidade pré-histórica onde indivíduos se comunicam por meio de uma linguagem ryliana. Tal linguagem repousa unicamente na relação do discurso intersubjetivo com os objetos públicos. Seu vocabulário é descritivo. Esses personagens fazem um uso primário dos operadores da lógica de predicados e incluem o modo verbal subjuntivo.

Em um estágio seguinte, a linguagem deles é enriquecida com recursos semânticos - como significado, referência, verdade e falsidade - permitindo-lhes expressar os significados de certas palavras ou proposições. Num segundo estágio do enriquecimento dessa linguagem, adiciona-se o discurso teórico, isto é, sobre não-observáveis, o que permite compreender e esclarecer o que é observável e formar um quadro geral de conhecimentos.

Eis que nasce um gênio nessa comunidade! Ele se chama Jones, é um precursor do behaviorismo metodológico⁴³. Jones desenvolve uma teoria para explicar o comportamento. Segundo a teoria, a fala pública é apenas o resultado de um processo que se inicia com episódios internos, compostos por um discurso interno que possuiria o mesmo modelo da fala pública, isto é, usaria as mesmas palavras e possuiria as mesmas regras semânticas. De acordo com a teoria de Jones, a causa do comportamento inteligente seria o discurso interno (SELLARS, 1956/1997, p. 103). Esse discurso, Jones batiza de “pensamento”. Então, pensamentos seriam não observáveis utilizados para explicar o comportamento, espécies de termos teóricos da teoria de Jones acerca da psicologia.

Jones, então, decide treinar seus conterrâneos para usarem essa teoria no intuito de interpretar o comportamento dos outros atribuindo-lhes estados internos. A partir daí, as pessoas ganham um recurso para explicar seus próprios comportamentos: a auto-atribuição de estados internos. No estágio seguinte, os indivíduos dessa comunidade não precisam mais observar seus próprios comportamentos para atribuírem pensamentos a si mesmos, o que os leva a falar do acesso privilegiado aos próprios estados mentais. Assim, o que começa como um vocabulário puramente teórico usado para a explicação do comportamento termina ganhando o papel de permitir ao sujeito falar sobre si mesmo e falar consigo mesmo. Mas os conceitos usados para isso são primária e essencialmente intersubjetivos. O acesso privilegiado, então, tem seu germe na intersubjetividade.

Inicialmente, os pensamentos surgem como termos não observáveis, e não como experiências internas. Importante ressaltar que, o fato de terem sido postulados como não observáveis não implica em que Jones precise supor que eles não existam. Jones usa a teoria para explicar os comportamentos como resultado de um processo que começa de dentro para fora, mas o mito de Jones, da maneira como é contado, é compatível com a ideia de que a

⁴³ Nesse ponto, Sellars diferencia o behaviorismo metodológico do analítico (ou filosófico). O behaviorismo metodológico está preocupado em construir uma ciência da psicologia. Não está preocupado em fazer uma análise lógica do discurso mentalista de senso comum nem nega o acesso privilegiado do sujeito aos próprios estados mentais. O behaviorismo analítico ou filosófico é, segundo ele, a tese extremamente austera segundo a qual os conceitos da psicologia behaviorista devem ser construídos a partir de um vocabulário pertencente ao comportamento aberto e aos demais objetos públicos. Porém, nem mesmo as ciências duras, como a física, têm esse tipo de restrição, elas usam um vocabulário teórico referindo-se a termos não observáveis. Para Sellars, nada haveria no behaviorismo metodológico que impedisse o uso do vocabulário sobre não-observáveis.

habilidade para desenvolver um discurso interno se adquire somente após a aquisição de um discurso aberto que é posteriormente interiorizado.

De acordo com Sellars, o mito de Jones concilia (1) a ideia de Skinner, Wittgenstein, Carnap e outros de que a linguagem é adquirida publicamente com (2) a ocorrência de episódios internos e privados. “Também deixa claro que essa privacidade não é uma ‘privacidade absoluta’” (SELLARS, 1956/1997, p. 107, aspas duplas do original substituídas por aspas simples, tradução nossa). O comportamento continua sendo evidência para a existência desses episódios internos, construídos a partir do que está publicamente disponível, em especial, a atividade linguística intersubjetiva⁴⁴.

O mito de Jones traz importantes elementos para o desenvolvimento do interpretivismo acerca do mental. Primeiro, mostra como os estados internos podem ser compreendidos a partir da atividade pública de comunicação. Mostra também que a própria noção de pensamento como algo privado pode ter sido construída apenas depois da construção da linguagem pública. Argumenta que a linguagem pública é necessária para que se possam atribuir pensamentos privados às pessoas. Além disso, essa atribuição de pensamentos pode ter sido feita, inicialmente, para dar sentido aos comportamentos públicos das outras pessoas, e só depois se voltar para algo feito pelo sujeito para si mesmo.

Sellars apresenta e defende a tese de que nós usamos um vocabulário teórico para explicar e prever o comportamento inteligente. Essa tese costuma ser chamada de “teoria-teoria”, a ideia de que nós usamos uma teoria que atribui estados mentais para previsão e explicação comportamental. Essa ideia certamente é importante para o desenvolvimento do interpretivismo. Porém, Dennett (1991a) compreende a atribuição de estados mentais como algo que o ser humano faz naturalmente, independentemente da linguagem. No mais, tanto Dennett quanto Davidson compreendem a atribuição de estados mentais como sujeita à indeterminação, de modo que não pode se sujeitar a teorias que imponham generalizações estritas como as das ciências naturais. Então, o termo “teoria” teria que ser usado em um sentido bem mais amplo para poder designar a atribuição de estados mentais em uma perspectiva interpretivista.

O interpretivismo compreende a atribuição de estados mentais como condição para que possamos dar sentido ao comportamento dos outros seres. No caso do interpretivismo de Davidson, os seres linguísticos. Dentro dessa abordagem, não há outra maneira de nos

⁴⁴ O mito de Jones não termina aqui. Ele continua com uma explicação para as impressões sensoriais. “Jones agora faz para as impressões sensoriais o que ele havia feito anteriormente para os pensamentos” (BRANDON, 1997, p. 177, tradução nossa). Como sói acontecer quando entramos no domínio das sensações, a explicação aqui ganha contornos bem mais complexos.

relacionarmos com os outros sem essa atribuição. O mito de Jones abre esse caminho, contando uma história a partir da qual se torna possível compreendermos os estados mentais primeiramente como algo que se atribui a partir de uma perspectiva externa, mas que não deixam, por isso, de ser estados internos. Só não podem ser objetos isolados perante o sujeito.

Conforme Dennett (1987), depois que as bases foram fundamentadas por Sellars e Quine, surgem duas posições acerca de quais são os princípios da interpretação. A diferença entre elas não é muito profunda, mas é uma diferença de ênfase. Uma é fundamentada em um princípio normativo segundo o qual nós atribuímos às pessoas as atitudes proposicionais que *devemos* atribuir em determinadas situações. A outra se baseia em um princípio projetivo, no qual as pessoas atribuem às outras as atitudes proposicionais que elas próprias teriam naquela circunstância.

Ele se coloca junto com Davidson e com qualquer autor que defenda alguma versão dos princípios de caridade e de racionalidade como partidário do princípio normativo. O princípio projetivo foi proposto para se contrastar ao princípio de caridade. A ideia é que nós nos projetamos no interlocutor e atribuímos a ele a atitude proposicional que nós teríamos frente à situação. Conforme Dennett, os dois tipos de princípio são perfeitamente compatíveis.

Essa compatibilidade pode ficar clara a partir de uma compreensão do interpretivismo de Davidson. De modo geral, pode-se dizer que os estados internos do sujeito se formam a partir de suas relações com os outros sujeitos e o mundo circundante. Por isso, tender-se-á a fazer uma atribuição de estados mentais aos outros de acordo com nossa própria percepção, mas essa percepção foi construída nessa relação, o que lhe dá certo grau de confiabilidade. Essa confiabilidade apoia também os princípios de caridade e de racionalidade.

Só é preciso tomar cuidado para não levar esses princípios ao pé da letra demais. Não é que nós estejamos sempre pensando as mesmas coisas que as outras pessoas, não é que todas as pessoas pensem igual. É que há um fundo de crenças muito básicas que são compartilhadas. É um fundo com muitas crenças, ele forma a base da nossa racionalidade, sem ele não conseguiríamos pensar e nos comunicar com as outras pessoas. Isso não impede, contudo, que cada pessoa tenha uma forma própria de pensar sobre as situações. Por isso, não existe *a* atribuição correta de atitudes proposicionais. Essa atribuição está sujeita à indeterminação.⁴⁵

⁴⁵ Ver seção 4.1.3.

1.6.1. Indeterminação da tradução

A noção de indeterminação da tradução, defendida por Quine, é aceita tanto por Dennett quanto por Davidson, mas de modos distintos. Isso possui consequências para a compreensão que cada um desses autores possui da ontologia das atribuições de estados mentais, conforme se tratará no último capítulo. De modo bem geral, Quine afirma que “manuais para traduzir uma língua em outra podem ser estabelecidos de maneiras divergentes, todas compatíveis com a totalidade das disposições verbais, porém incompatíveis entre si” (1960/2010, p. 51). Em outras palavras, Quine afirma ser possível haver manuais de tradução incompatíveis e igualmente corretos para a mesma língua. Conforme Hylton (2018), isso é uma consequência da ideia quineana de que o critério para uma comunicação bem sucedida é a interação fluente. Tudo o que se requer de um método de tradução é que ele nos permita a comunicação com os falantes de outra língua, nada impede que isso seja feito de diferentes formas.

Quine defende que não há ideias pré-linguísticas sendo comunicadas em forma de linguagem, mas o que há é a interação linguística. Em outras palavras, não há uma base fixa de significado sobre a qual os diferentes idiomas seriam construídos. Em geral, temos uma tendência intuitiva a pensar que, para cada frase de um idioma, há uma correspondente em outro idioma, mas Quine rejeita isso.

Heil (1998, p. 181) explica que a doutrina da indeterminação da tradução não diz respeito apenas a uma dificuldade ou mesmo impossibilidade epistemológica de escolher um manual de tradução em detrimento do outro. Ela traz consigo a tese ontológica de que a “verdade” está apenas na construção de um manual de tradução adequado a todas as evidências disponíveis. Se dois manuais divergentes conseguirem satisfazer esses critérios, então não existirá um que seja verdadeiro frente ao outro.

De acordo com Miller (2010), o argumento de Quine segue em um sentido epistemológico para tirar conclusões ontológicas. A ideia que ele pretende defender é que não há um fator objetivo capaz de determinar que um manual de tradução seja correto e o outro não seja. “Quando nos defrontamos com dois tipos diferentes de manuais de tradução, não ocorre de podermos ‘estar certos’ em nossa escolha: alguns manuais de tradução seriam mais úteis, mais elegantes, mais naturais, mais simples etc., mas essas seriam considerações puramente pragmáticas” (MILLER, 2010, p. 145, aspas duplas do original substituídas por aspas simples, tradução nossa). O que determina que uma frase de um idioma seja tradução de outro idioma são fatores comportamentais. Nesse sentido, uma frase traduzida tem, sim, o

mesmo significado que teria no idioma original, mas não há nenhum fato mais profundo acerca do significado que possa fixar a tradução correta.

Isso está de acordo com a noção de verdade defendida pelo próprio Quine, para quem essa noção só faz sentido dentro de um corpo teórico que inclui uma realidade, em certa medida, postulada. Ou seja, dentro de um corpo de conhecimentos, algumas frases parecem estar bem estabelecidas em confronto com a realidade, enquanto outras são construtos teóricos que bem poderiam ser postulados de outra forma, e igualmente compatíveis com aquelas frases bem estabelecidas. A partir desses postulados, toda uma visão de mundo é estruturada, sendo que ela pode sofrer ajustes pontuais. Isso inclui também os significados dos termos. Porém, visões de mundo completamente diferentes também seriam possíveis.

É quando nós nos voltamos para o meio de uma teoria realmente presente, pelo menos hipoteticamente aceita, que nós podemos, e de fato falamos, ponderadamente, desta e daquela frase como verdadeira. Faz sentido aplicar “verdadeiro” a uma frase expressa em termos de uma teoria dada e vista desde o interior da teoria, acrescida de sua realidade postulada. [...] Que os enunciados são acerca de entidades postuladas, são significantes somente em relação a um corpo de teoria circundante, e são justificáveis somente pela suplementação da observação por método científico, não importa mais; pois as atribuições de verdade são feitas do ponto de vista do mesmo corpo de teoria circundante, e estão no mesmo barco (QUINE, 1960/2010, p. 49, aspas do original).

Isso, contudo, não leva Quine a se satisfazer com uma perspectiva relativista, pois dentro de um corpo de crenças, “continuamos levando a sério nosso próprio agregado particular de ciência, nossa própria teoria de mundo particular ou a frouxa trama total de quase-teorias” (QUINE, 1960/2010, p. 49). Assim, por mais que a noção de verdade passe por ajustes constantes, ela continua fazendo parte de um corpo de crenças contra o qual os ajustes são feitos. Esse corpo de crenças se ajusta também a uma série de fatos. Nesse sentido, então, há mais do que uma forma pela qual a verdade pode se ajustar à realidade. Não há unicidade, mas também não há relativismo.

Para compreender porque Quine defende essa ideia, é preciso resgatar a noção de tradução radical. Quine (1960/2010, p.53) propõe o pensar sobre a “tradução de uma língua de um povo ainda não tocado”, pois isso traz luz para as questões acerca da linguagem. Para que se possa avaliar como isso ocorre, as evidências disponíveis ao tradutor não incluirão intérpretes bilíngues, nenhum conhecimento prévio da língua a ser traduzida e nenhum dicionário. O tradutor estará em uma situação de completa ignorância também acerca dos aspectos da história ou hábitos culturais daquele povo cuja língua pretende compreender. Ele deverá se basear apenas no contexto presente a si e ao falante e naquilo que se apresenta perante os seus sentidos.

No processo da tradução radical, os sons emitidos pelo falante serão anotados pelo tradutor juntamente com observações sobre o contexto da fala. Por exemplo, “um coelho passa correndo, o nativo diz ‘Gavagai’ e o linguista anota a frase ‘Coelho’ (ou ‘Olha, um coelho’) como uma tradução provisória, sujeita a ser testada em casos ulteriores” (QUINE, 1960/2010, p.53, aspas duplas do original substituídas por aspas simples).

Para Quine, os estímulos sensoriais desempenham um papel relevante, pois o termo “Gavagai” pode ter outros significados além de “coelho”, por exemplo, se o coelho for branco, ou se o nativo quer se referir a animal em geral. Por isso, a hipótese de que “Gavagai” quer dizer “coelho” precisará ser testada. Nesse intuito, o tradutor tomará notas das diferentes situações de emissão do som “Gavagai”. Outra tarefa do tradutor é descobrir os diversos comportamentos do falante associados a concordância ou discordância ou a afirmação e negação. Os comportamentos associados a essas atitudes podem ser diferentes em diferentes culturas. Portanto, aqui também será importante a realização de experimentos repetitivos e anotações que associam os sons emitidos pelo falante ao contexto da fala. A comunicação desempenhará um papel importante, pois o tradutor irá imitar os sons do falante diante de determinados objetos do mundo e observar o comportamento de seu interlocutor de modo a descobrir o que é, para ele, afirmação e o que é negação. Quine, assim, traz o significado dos termos como associado ao estímulo sensorial.

Porém, ocorre que, com apenas essas evidências disponíveis, Quine argumenta que não é possível determinar, por exemplo, se “Gavagai” se refere a coelho ou a uma parte inseparável dos coelhos. As evidências disponíveis não permitem a solução dessa questão. Se preferimos a tradução “coelho”, é apenas por motivos pragmáticos e estilísticos, mas não há nenhum fato definitivo capaz de determinar a tradução que seria correta.

Miller (2010, p. 159 e segs.) explica que há outro argumento que Quine considera ainda mais consistente na defesa da indeterminação da tradução. Ele compara a indeterminação da tradução à subdeterminação das teorias físicas. Se as teorias físicas são, em algum nível, subdeterminadas pelos dados sensoriais, então a tradução dessas teorias de um idioma para outro também sofrerá de subdeterminação. A subdeterminação é a tese de que há diversas teorias incompatíveis entre si, mas compatíveis com o mesmo conjunto de dados observáveis. Aplica-se, em geral, as teorias acerca do que é inobservável, mas Quine a estende ao discurso de senso comum sobre as coisas visíveis (QUINE, 1970, p. 183).

Conforme Quine (1970, p. 179), “teorias físicas podem discordar umas das outras e ainda ser compatíveis com toda informação possível mesmo no sentido mais amplo. Numa palavra, elas podem ser logicamente incompatíveis e empiricamente equivalentes”. A

subdeterminação ocorre porque os termos teóricos da física são construídos sobre as evidências observacionais, mas nada impede que teorias diferentes sejam compatíveis com as mesmas evidências, ainda que incompatíveis entre si.

Quine entende que nossa visão de mundo compõe uma rede de crenças que enfrenta a experiência não de forma individual, ou seja, não é possível contrastar as nossas proposições diretamente com a realidade uma a uma para dizer se elas são verdadeiras ou falsas, pois nessa rede umas proposições estão amarradas a outras, por vezes com laços frouxos que podem facilmente ser soltos e refeitos, mas por vezes com nós muito firmes. Essa é a noção de *holismo epistemológico*. É um corpo completo de enunciados que enfrenta o juízo experimental, e não os enunciados singulares.

Quando ocorre a divergência entre uma proposição e a realidade, então, uma série de ajustes é feita na rede, algumas proposições serão alteradas, outras cairão e algumas serão mantidas intactas. Quais proposições serão alteradas, quais cairão e quais serão mantidas? Na concepção de Quine, há diversas formas de fazer esses ajustes, todas igualmente apropriadas frente às evidências observacionais. Isso se aplica às teorias da física e é por isso que elas são indeterminadas. Mas, para Quine, isso se aplica também aos nossos corpos cotidianos de conhecimento. Essa é uma das razões pelas quais o autor defende que isso se aplica às traduções. Ou seja, assim como as teorias da física são subdeterminadas, para Quine, as traduções são indeterminadas.

A noção de indeterminação da tradução proposta por Quine é relevante para a compreensão que ele tem da noção de significado. Para que se determine o significado de um termo, uma convenção é formada dentro de uma comunidade linguística. Ora, as convenções se constituem a partir de relações comportamentais entre as pessoas envolvidas. Assim, uma vez que as linguagens se constroem a partir dessas convenções, o contexto da tradução radical se torna bastante esclarecedor na compreensão do que seja o significado e como se forma a linguagem.

Conforme Miller (2010, p. 149), o argumento de Quine vai além de um ceticismo epistemológico, embora parta daí, ele se aprofunda para a tese ontológica de que não há significados subjacentes capazes de fixar a tradução correta. Apresentado pelo viés epistemológico, o argumento seria que, diante de dois divergentes manuais de tradução compatíveis com os estímulos sensoriais relevantes disponíveis, não teríamos como *saber* - qual seria o correto. Ocorre que não há o manual correto. Na verdade, ambos os manuais estão corretos, e nada impede que haja outras opções de manuais corretos disponíveis. Nesse

sentido, o argumento epistemológico é usado para trazer à tona a tese ontológica de que não existem significados fixos subjacentes à tradução.

Quine é um autor extremamente influente no que estaria por vir dentro da filosofia analítica. A noção de indeterminação da tradução será bastante importante para podermos pensar sobre o status ontológico que Davidson e Dennett, ambos alunos de Quine, dão às atribuições de estados intencionais. Para adiantar, a tese da indeterminação da tradução não conduz Quine nem seus alunos à adoção de uma postura antirrealista.

2. INTERPRETIVISMO DE DONALD DAVIDSON: COMO O DISCURSO CONSTITUI UMA REDE DE ESTADOS MENTAIS

A interpretação é uma atividade com que os seres humanos estão quotidianamente engajados, algo que fazemos sem nenhuma dificuldade. Porém, segundo Davidson, não é fácil explicá-la, pois ela depende de uma série de atribuições de estados mentais aos outros, sendo que a mente alheia ainda parece um mistério, uma vez que não temos acesso livre aos pensamentos das outras pessoas. Porém, deve-se levar em conta o resultado final: na maior parte das vezes, a comunicação entre as pessoas obtém sucesso. Por exemplo, alguém chega à biblioteca com seu notebook descarregado e não encontra uma tomada disponível. Então, percebe outra pessoa com seu computador ligado a uma extensão que possui mais entradas desocupadas. Silenciosamente, a pessoa que precisa ligar o computador olha para a outra e o olhar é retribuído. Ela mostra à outra o plugue de seu notebook e a dona da extensão acena positivamente com a cabeça. A pessoa liga seu notebook à extensão, faz um gesto de positivo com a mão e sorri para a outra. A partir desse momento, ambas voltam a seus trabalhos sem nenhum problema.

Naquela circunstância, as duas evitaram trocar palavras para não importunar os outros usuários da biblioteca, mas conseguiram se comunicar tranquilamente por meio da atribuição de crenças e desejos uma à outra. Isso ocorreu de modo instantâneo, sem que precisassem fazer nenhum tipo de elaboração teórica complexa. Nesse caso, ao que tudo indica, as interpretações que cada uma fez das crenças e desejos da outra foram bem-sucedidas e obtiveram resultados satisfatórios. Esse exemplo quer chamar a atenção para o fato de que a interpretação funciona, na maior parte das vezes, de modo fluido e natural. Nem sempre é assim, há erros de comunicação, discórdias ou mesmo situações em que o acordo ocorre por acaso. Porém, na maior parte das vezes nós somos capazes de compreender o que os outros dizem e de esclarecer o que queremos dizer com nossas palavras. Dentro do pensamento de Davidson, as respostas à questão da interpretação são muito centrais e podem iluminar outras questões filosóficas, especialmente acerca do mental e do significado.

De acordo com Miguens (2006, p.104):

[...] o conceito de crença tem para Davidson duas funções: (i) aquela que mais frequentemente se refere, que é permitir interpretar o comportamento de outra criatura, (ii) uma menos frequentemente sublinhada, que é ser o veículo para o sujeito capturar o conceito de verdade objectiva.

Para que possamos fazer juízos sobre o mundo, segundo Davidson, é preciso haver a linguagem e é a partir da linguagem que as mentes dos seres humanos são constituídas. Assim, ao compreender o mental linguisticamente, é possível compreendê-lo também como algo que se constitui a partir da relação com o mundo circundante, isso evita os problemas filosóficos oriundos de uma percepção do sujeito como algo isolado da realidade. Contudo, a posição de Davidson parece ter a desvantagem de ser uma abordagem parcial e, talvez, superficial do mental, pois ele considera apenas os aspectos linguísticos, mas a mente parece ser constituída de outras características cognitivas além das atitudes proposicionais.

Pensar na mente humana como algo que resulta de uma história evolutiva pode ser interessante, pois a relação com o ambiente molda a evolução dos seres vivos em diversos aspectos. No decurso da evolução, o ambiente permite que algumas criaturas se reproduzam e outras não. Assim como a linguagem se desenvolve no contexto da relação direta das criaturas entre si e com o mundo, o mesmo ocorre com todos os nossos traços cognitivos e com os de qualquer ser vivo. Portanto, ao mesmo tempo em que esse tipo de abordagem pode corroborar com as teses de Davidson acerca do mental, pode ir além, ajudando-nos a compreender os aspectos não linguísticos de nossas próprias mentes e dos demais seres vivos.

Este capítulo está estruturado em três seções, as quais possuem respectivas subseções. A primeira seção aborda a noção de interpretação radical. Para compreendermos a interpretação, é relevante compreender também a noção que Davidson possui de verdade para as linguagens naturais, além da noção de tradução de Quine.

A segunda seção está focada no princípio de caridade. Também trata da crítica de Davidson ao dualismo entre esquema conceitual e conteúdo perceptivo, chamado por ele de “terceiro dogma do empirismo”. Nessa seção, abordaremos ainda um debate sobre o realismo ou relativismo em Davidson.

A terceira e última seção será sobre a noção de triangulação. Nela, está incluída uma espécie de relatório de leitura de um dos livros mais recentes do psicólogo do desenvolvimento Michael Tomasello (2014), pois vemos nele um suporte empírico para as teses davidsonianas da triangulação e do princípio de caridade. Embora Tomasello concorde com Davidson que há um tipo de mente especificamente humana e que esse tipo de mente está linguisticamente estruturado, o psicólogo defende que existe pensamento pré-linguístico, sem o qual a linguagem convencional não poderia ter surgido na história evolutiva. Por um lado, podemos tomar isso como uma diferença terminológica entre os dois. Porém, Davidson parece quase que ignorar completamente os aspectos não linguísticos do mental e

provavelmente essa lacuna pode ser minimizada, se não completamente eliminada a partir de uma abordagem evolucionista mais detalhada como a de Tomasello.

2.1. Interpretação radical

Conforme Caro (1999), a noção de interpretação inicialmente é tida por Davidson como um estudo promissor para a filosofia da linguagem, mas aos poucos se torna tema central e integrador de todo o pensamento davidsoniano, juntamente com a noção de causa. A própria possibilidade de comunicação e o pensamento se ancoram na interpretação. Para decifrar um discurso, o intérprete precisará identificar os estados mentais do falante a partir da relação entre o mundo e aquilo que é dito, de modo que a mente do sujeito pode, assim, se revelar ao interlocutor. A noção de interpretação também é importante na compreensão do externismo de Davidson e de como ele constrói pontes entre subjetividade e objetividade.

A noção de interpretação radical se torna evidente a partir da possibilidade existente a um estrangeiro que desconhece completamente uma língua descobrir gradualmente o que significam os termos nessa língua. Por exemplo, “Kurt profere as palavras ‘Es regnet’ e sob as condições corretas nós sabemos que ele disse que está chovendo” (DAVIDSON, 1973/1984b, p. 125, tradução nossa). Como é possível descobrirmos isso? A possibilidade de atribuição de significado a um discurso enunciado em uma língua completamente desconhecida permite a compreensão de como atribuímos significado em geral, inclusive em nossa própria língua.

A tradução radical de Quine é muito importante para Davidson desenvolver sua noção de interpretação radical. Porém, Davidson (1984b, p.126) utiliza o trabalho de Quine para lidar com questões diferentes e por isso sua preocupação com a interpretação, ao invés da tradução. Enquanto Quine está preocupado com questões de epistemologia, Davidson está interessado nas questões do significado e dos estados mentais. Isso gera algumas diferenças entre as ideias dos dois autores.

Malpas (2015) explica algumas diferenças entre a noção de tradução radical de Quine e a interpretação radical de Davidson. Se em situações comuns a tradução é realizada a partir de algum conhecimento prévio da língua a ser traduzida, na tradução radical proposta por Quine, como já abordado, não há esse recurso. O trabalho é feito apenas com base no comportamento verbal e não verbal do falante e os estímulos sensoriais⁴⁶ que dão suporte a esse comportamento. Davidson, por sua vez, amplia um pouco a evidência disponível ao

⁴⁶ De acordo com Davidson (1995/2005c), Quine (1992) muda de posição e passa a considerar os eventos compartilhados entre falante e intérprete como importantes para o processo de interpretação.

intérprete, sendo-lhe permitido atribuir ao falante alguma atitude mental, como a intenção de sustentar determinadas frases como verdadeiras (na verdade, essa é uma condição *sine qua non* da interpretação). Também minimiza o papel dos estímulos sensoriais como evidências. Este último ponto representa uma oposição a Quine, pois Davidson considera o empirismo uma forma de subjetivismo.

Davidson usou o termo “interpretação radical” para qualquer interpretação de um discurso a partir de uma posição epistêmica que tenha apenas o mínimo de informações necessárias para que a interpretação possa ocorrer. Essa pode ser, por exemplo, a posição de uma criança ao aprender sua primeira língua (nesse sentido, todos seríamos intérpretes radicais) ou do linguista que precisa compreender um discurso proferido em uma língua completamente desconhecida. O termo é usado para se referir a uma interpretação feita a partir de uma base de dados limitada, específica e publicamente disponível (DAVIDSON, 1994a). Essa base de dados consiste no estabelecimento de relações entre aquilo que, na concepção do intérprete, o falante sustenta como verdadeiro ou falso e o mundo compartilhado no momento em que o discurso é proferido. Diante apenas desses dados, o intérprete radical é alguém capaz de, por exemplo, construir regras para interpretar um discurso estrangeiro.

Todos nós estamos quotidianamente envolvidos com a tarefa da interpretação, somos intérpretes reais, isso é pressuposto para a compreensão de um discurso. A construção do intérprete radical foca nas situações mais extremas, ou seja, quando é necessário compreender um discurso enunciado em uma língua desconhecida. O intuito é didático: isso pode trazer luz para os processos quotidianos de interpretação que, por sua vez, ajudam a compreender tanto a questão do significado quanto os processos mentais do falante. Nesse sentido, “[t]oda compreensão do discurso de outrem envolve interpretação radical” (DAVIDSON, 1973/1984b, p. 125, tradução nossa)⁴⁷. Ocorre que nas nossas interações quotidianas, temos uma série de informações linguísticas compartilhadas com o interlocutor previamente ao processo de comunicação. O intérprete radical se caracteriza por ter essas informações reduzidas ao mínimo.

⁴⁷ Posteriormente, em resposta a uma crítica de Fodor e Lepore (1994) a essa ideia, Davidson (1994a) esclarece que não há uma identificação entre o intérprete radical e o intérprete real, pois o intérprete real não precisa ter apenas as evidências disponíveis ao intérprete radical para a realização de seu trabalho. Quer dizer, embora a interpretação quotidiana tenha elementos da interpretação radical, as duas não são a mesma coisa. A interpretação radical é uma espécie de idealização que ajuda na compreensão da interpretação real, assim como os experimentos de pensamento filosóficos em geral são idealizações que permitem o esclarecimento de alguma questão.

É importante pontuar que a interpretação radical não pode tomar como evidência as intenções e crenças detalhadas do falante, pois esse é um aspecto a ser esclarecido conjuntamente à interpretação. Há uma interdependência entre crença e significado: alguém afirma uma frase como verdadeira em decorrência do que crê que ela significa e do que acredita ser verdade acerca do mundo. Ao compreender o significado de um discurso as intenções e crenças do falante serão simultaneamente esclarecidas, como parte do projeto de interpretar. Portanto, tomá-las como dado inicial seria petição de princípio (DAVIDSON, 1973/1984b, p.127). Outro critério é que a evidência disponível ao intérprete deve ser apresentada sem o uso de conceitos linguísticos como “significado, interpretação e sinonímia” (1973/1984b, p.128).

Convém observar que estamos tratando aqui de crenças que subjazem a qualquer frase, seja ela declarativa ou não. Por exemplo, se alguém faz a pergunta “Que horas são?”, para que seja possível compreender o que se está perguntando, é preciso que falante e intérprete compartilhem de certas crenças, por exemplo, acerca do significado das palavras que compõem essa pergunta, acerca de que se trata de uma pergunta (não de uma exclamação nem de uma ordem), de que ambos estão falando do momento que é presente a ambos, de que estão lidando com o mesmo sistema de medição do tempo, e assim sucessivamente. As crenças compartilhadas subjacentes à compreensão dessa pergunta são tantas que, ainda que fizéssemos uma lista longa, correríamos um sério risco de deixar algumas de fora. Portanto, o vínculo ente crença e significado não é algo que isola os as frases declarativas e seus significados. Ao contrário, extrapola as frases declarativas e aquilo que está sob alvo da interpretação, atinge todo tipo de discurso e, provavelmente qualquer comportamento intencional.

Em outras palavras, na interpretação radical, o intérprete não pode ter de antemão dados detalhados sobre o que significam os sons proferidos pelo falante nem sobre suas atitudes proposicionais, primeiro porque o trabalho do intérprete é justamente descobrir os significados daquilo que está sendo proferido pelo falante. Segundo, porque um conhecimento detalhado das atitudes proposicionais do falante não é algo que esteja disponível publicamente para nenhum intérprete. Os significados e as atitudes proposicionais do falante são compreendidos ao longo do processo de interpretação por serem interdependentes. “O que se busca, então, é uma abordagem que produza uma interpretação de uma das palavras do falante ao mesmo tempo que proporcione uma base para a atribuição de crenças e desejos ao falante. Tal abordagem visa prover uma base para, em vez de assumir, a individuação das atitudes proposicionais.” (DAVIDSON, 1990/2002a, p. 95). Assim como ocorre na interpretação

radical, ocorre nos processos de comunicação reais e quotidianos: pela interpretação da fala de nossos interlocutores, sabemos o que aquela pessoa pensa, deseja ou crê (DAVIDSON, 1990/2002a, p. 97).

Embora não seja permitido ao intérprete radical um conhecimento prévio detalhado acerca do que significam as palavras proferidas na língua estrangeira a ser interpretada, está pressuposto um compartilhamento aproximado de uma série de conceitos. Um deles, provavelmente o mais importante na construção da teoria de Davidson, é o conceito de *verdade*. Mas também é preciso um acordo aproximado acerca das noções de intenção, crença, desejo, afirmação e negação (entre outros conceitos), noções amplas sobre como as pessoas se comportam em diversas circunstâncias, padrões humanos de discriminação dos objetos do mundo, e assim sucessivamente. A base de dados compartilhada entre falante e intérprete é enorme.

O que o intérprete não pode saber? Ele não é um telepata, não consegue fazer uma leitura mágica e detalhada dos pensamentos do interlocutor. Ele não tem dicionários, não tem um tradutor bilíngue a seu lado. Também não conhece nenhuma palavra do idioma a ser interpretado. Não tem, portanto, nenhum acesso ao que significam as palavras do interlocutor, não tem um livro explicando o que se passa na mente alheia. Contudo, há uma base de dados disponível para o intérprete, muitos desses dados não são linguísticos. Ele precisará construir gradualmente uma compreensão da linguagem do interlocutor, da mesma forma como aprendeu os significados dos termos em sua própria língua. Nesse sentido, sua interpretação será construída “do nada” [“from scratch”] (DAVIDSON, 1994a, p.125).

Assim, para que se dê a interpretação radical, não há nenhum tipo de acesso privilegiado à mente alheia além do acesso que normalmente se tem nos processos de comunicação quotidianos. Sempre que nos engajamos com a linguagem, um processo de interpretação entra em cena e, se aceitamos os argumentos de Davidson, entendemos que, antes de começar a conversa, já temos um conhecimento bastante amplo do que se passa na mente de nosso interlocutor. Quanto mais próximo esse interlocutor estiver de nosso entorno linguístico, mais saberemos acerca do que se passa em sua mente. Por exemplo, eu tenho um conhecimento prévio maior da mente de minha amiga de infância do que da mente de um colega de curso. Porém, sei mais da mente do colega que do indiano que é professor de yoga. De todo modo, por mais que o interlocutor esteja culturalmente afastado de minha história de vida, eu compartilho com ele de um fundo enorme de noções a partir do qual é possível a comunicação.

Durante o diálogo, a mente do outro se abrirá gradualmente diante do interlocutor, revelando suas crenças, desejos, medos, esperanças... Qualquer coisa que se passe naquela mente e que possa ser linguisticamente expressa poderá se desvelar perante o interlocutor. Há, inclusive, coisas que o falante pode não querer expressar, mas que se manifestam por meio de alterações no tom da voz ou em modos diferentes de olhar, por exemplo. Há aquilo que comunicamos, isto é, algo que queremos dizer e há também o que manifestamos sem querer⁴⁸.

Importante notar que isso não significa um acesso privilegiado de terceira pessoa à mente do falante. Davidson (2001a) concorda que há uma assimetria entre o acesso que o sujeito tem aos seus estados mentais e o acesso que temos aos estados mentais alheios. Concorda, ainda, que o acesso de primeira pessoa é mais nítido e amplo. Porém, esse é o processo natural de comunicação: uma vez que a mente se constrói linguisticamente e que a linguagem é pública, a mente se revela publicamente por meio da linguagem. É isso que dá à concepção de Davidson acerca do mental o epíteto de interpretivismo.

Para que possamos considerar a interpretação radical bem-sucedida, é necessário que o intérprete compreenda o significado de uma frase proferida pelo estrangeiro. Contudo, quando uma frase é proferida, ela é uma entre infinitas frases que poderiam ser emitidas (1973/1984b, p. 128). Uma teoria geral de como se dá a interpretação deverá, portanto, se aplicar a uma quantidade de enunciados potencialmente infinita.

Aqui, estamos diante de um problema: como nós, seres finitos, somos capazes de utilizar e compreender uma linguagem na qual infinitas frases podem ser construídas? Para compreendermos uma oração, é preciso compreendermos o significado de termos gerais, termos singulares e predicados, além das regras de concatenação entre esses elementos. Para apreendermos cada item do vocabulário e cada regra gramatical, demoramos um tempo finito (DAVIDSON, 1965/1984a, p. 5). Por isso, Davidson busca uma teoria que permita compreender o que significam todas as frases de uma língua e conclui que a definição de verdade de Tarski é um método promissor para determinar o significado de cada uma.

2.1.1. Davidson e a verdade nas linguagens naturais

Tarski (1936) empenha-se em definir a verdade para as linguagens formais (GÓMEZ-TORRENTE, 2015), considerando que, embora a noção de verdade nos seja

⁴⁸ Aparentemente, esse não é um ponto desenvolvido por Davidson, mas para os fins dessa tese convém notar que o intérprete atento pode chegar a compreender aspectos da mente de seu interlocutor que este não percebe acerca de si. Pode, inclusive, alertar seu interlocutor acerca dos estados mentais dele (o interlocutor).

familiar e intuitiva, ela também é fugidia quando tentamos defini-la para as linguagens naturais, pois gera ambiguidades e paradoxos⁴⁹. O predicado de verdade para uma linguagem-objeto é apresentado por Tarski como pertencendo a uma metalinguagem⁵⁰. A construção apropriada de uma definição de verdade, para ele, deverá refletir uma concepção intuitiva de verdade e se aplicar às frases que são intuitivamente verdadeiras na linguagem objeto. Para uma teoria da verdade ser apropriada, Tarski estabelece como critério a convenção-V⁵¹ segundo a qual todas as frases declarativas de uma linguagem formal devem ser derivadas do seguinte esquema de frase: “*S* é verdadeira se, e somente se, *p*”, onde *S* deve ser substituída pelo nome de uma frase na linguagem-objeto e *p* é a tradução dessa frase em metalinguagem, gerando as chamadas “frases-V” (cada frase-V estabelece as condições de verdade de uma frase da linguagem-objeto). Por exemplo: “Chove” é verdadeira se, e somente se, chove. O que está entre aspas deve ser considerado como o nome de uma frase que ocorre na linguagem-objeto. Então, a convenção-V determina que uma teoria da verdade é satisfatória se gera uma frase-V para cada frase da linguagem-objeto.

Embora Tarski estivesse convencido de que seu trabalho sobre a verdade não se aplicaria às linguagens naturais, Davidson o considera bastante promissor, desde que algumas alterações sejam feitas para acomodar os problemas que surgem na passagem das linguagens formais para as naturais. Um dos problemas é incorporar à teoria da verdade a presença de indexicais⁵² (são o que ele chama de “aspectos demonstrativos” das linguagens naturais). Ocorre que, nas linguagens naturais, o momento do proferimento e a pessoa que fala são

⁴⁹ O principal problema, aqui, é o dos paradoxos semânticos. De acordo com Santos (2014), há duas características desses paradoxos: (1) envolvem conceitos semânticos como verdade, referência, definição, satisfação etc. e (2) como acontece aos paradoxos em geral, possuem uma conclusão contraditória ou absurda. Eubúlides de Mileto (séc.IV a.C.) é tido como o primeiro proponente do paradoxo do mentiroso, o mais famoso dos paradoxos semânticos. Consiste na seguinte afirmação de um sujeito: “Estou mentindo”. Ora, se ele estiver dizendo algo verdadeiro, estará mentindo e, se estiver dizendo algo falso, estará falando a verdade. Santos explica que Tarski se deparou com os paradoxos semânticos como um obstáculo na busca de uma definição de verdade. O problema surge quando uma linguagem possui o seu próprio predicado de verdade e pode ser superado se o predicado de verdade for apresentado como pertencendo a uma metalinguagem (ver a nota de rodapé seguinte), criando uma hierarquia das linguagens. Para uma linguagem *L* qualquer, o conceito de frase verdadeira em *L* deve, portanto, ser construído em uma metalinguagem. A verdade, assim, é um predicado monádico de uma metalinguagem. Por isso, Tarski só se propõe a definir a verdade para as linguagens formais: as linguagens naturais já contêm o predicado de verdade, o que torna inevitáveis os paradoxos semânticos. Por isso, qualquer teoria semântica das linguagens naturais, seguindo o método de Tarski, seria inconsistente.

⁵⁰ A linguagem-objeto é aquela sobre a qual se fala. Metalinguagem é a linguagem usada para falar sobre a linguagem-objeto. Em outras palavras, metalinguagem é uma linguagem cujos termos se referem a outra linguagem.

⁵¹ Também conhecida como “convenção-T”, optamos por “convenção-V”, pois a letra “T”, aqui, se refere a “truth”, isto é, “verdade” em nossa língua.

⁵² Indexicais são expressões cuja referência depende do contexto. Conforme Ruffino (2014, p.1), os indexicais incluem pronomes pessoais (“eu”, “ele”, “nós”, etc.), pronomes demonstrativos (“isto”, “aquilo”, etc.), advérbios (“aqui”, “agora”, “amanhã”, etc.) e todo tipo de expressão cuja referência só possa ser determinada dentro de uma situação específica. A linguagem natural se caracteriza por estar permeada pelos termos indexicais.

essenciais para que se avaliem as condições de verdade de uma frase. Em outras palavras, a verdade em uma linguagem natural deve ser caracterizada relativamente a pelo menos esses marcadores: tempo e falante (1973/1984b; p. 131; 1975/1984d, p. 149)⁵³, gerando frases-V como a seguinte: “Para todos os falantes de alemão x e todos os instantes t , ‘Das ist Weiss’ é verdadeira falada por x em t se, e somente se, o objeto mostrado por x em t é branco” (1975/1984d, p. 151; aspas duplas do original substituídas por aspas simples, tradução nossa).

A interpretação de predicados e nomes comuns depende fortemente de elementos de indexação na fala, tais como demonstrativos e tempos verbais, uma vez que são esses elementos que mais diretamente permitem aos predicados serem ligados a objetos e eventos no mundo. (Para acomodar elementos de indexação, as teorias da verdade do tipo das propostas por Tarski devem ser ampliadas [...]) (DAVIDSON, 1990/2002a, p. 100).

Uma das vantagens da ideia de Tarski, para Davidson, é que ela produz as condições de verdade de qualquer frase de uma linguagem sem precisar recorrer a nada mais do que à própria frase, ou seja, é ontologicamente leve. “A menos que a frase original mencione mundos possíveis, entidades intensionais, propriedades ou proposições, a declaração de suas condições de verdade também não mencionará” (DAVIDSON, 1973/1984b, p. 132, tradução nossa). Contudo, ao aplicar a ideia de Tarski para as linguagens naturais, Davidson não acredita na possibilidade de definir a verdade, ele trata o predicado de verdade como primitivo. As preocupações de Davidson e Tarski, portanto, são diferentes. Enquanto Tarski busca uma definição de verdade para as linguagens formais, Davidson está em busca de uma regra finita a partir da qual seja possível explicar a interpretação de infinitas frases em uma linguagem natural e encontra em Tarski o que procura.

O interessante da concepção de Tarski, para Davidson, é que ela provê uma espécie de pilar no qual uma abordagem da tradução e da interpretação pode se apoiar. Se uma teoria da verdade implica para cada frase da linguagem-objeto uma frase-V, isso serve de base para a tradução. Por exemplo, “Es regnet” é verdadeira se, e somente, se, chove.

Por um lado, nós temos frases-V na forma:

(T) ‘Es regnet’ é verdadeira-em-alemão quando dita por x no instante t se, e somente se, chove próximo a x em t .

Por outro lado, nós temos a evidência na forma:

(E) Kurt pertence à comunidade de fala germânica e Kurt sustenta como verdadeiro ‘Es regnet’ ao meio dia e chove próximo a Kurt ao meio dia. (DAVIDSON, 1973/1984b, p. 135, tradução nossa)

⁵³ De acordo com Davidson (1975/1984d, p. 150), é possível que haja outros marcadores também.

Assim, a tradução pode construir uma frase-V na metalinguagem para cada frase a ser traduzida da linguagem-objeto. Durante o processo de interpretação, a teoria da verdade pode ser testada e reformulada sempre que necessário, com base nas evidências disponíveis ao intérprete. Ora, mas Kurt poderia estar enganado acerca de se chove próximo a ele. O método para a interpretação, segundo Davidson, consiste em maximizar a concordância, tão frequentemente quanto possível, acomodando os diversos tipos de engano⁵⁴.

2.1.2. *É possível definir a verdade?*

Um problema que gera uma série de desdobramentos complicados, segundo Davidson (1996/2002b), é que quando os filósofos se debruçam sobre um conceito, tomam outros igualmente problemáticos como básicos. Porém, muitos dos conceitos filosóficos são fundamentais e deveriam ser estudados não por definições isoladas, mas a partir das relações entre eles.

No geral, os conceitos para os quais os filósofos chamam atenção, como verdade, conhecimento, crença, ação, causa, o bom e o certo, são os mais elementares conceitos que temos, conceitos sem os quais (estou inclinado a dizer) não teríamos absolutamente quaisquer conceitos. Por que então deveríamos esperar ser capazes de reduzir estes conceitos deflacionariamente a outros conceitos que são mais simples, claros e básicos? (DAVIDSON, 1996/2002b, p. 111).

Dentre esses conceitos básicos, de acordo com Davidson, está o de verdade. Para que algo interessante possa ser dito sobre ele, é preciso que o relacionemos a outros conceitos, de preferência que sejam igualmente básicos.

Davidson, assim como Tarski, demonstra uma simpatia pela caracterização aristotélica de verdade: “Dizer do que é que não é, ou o que não é que é, é falso, enquanto dizer do que é que é, ou do que não é que não é, é verdade” (DAVIDSON, 1996/2002b, p. 112)⁵⁵. A vantagem dessa forma de caracterizar a verdade, como vimos, é que não acrescenta a ela conceitos ulteriores como “correspondência à realidade”; “estados de coisas”; “fatos” etc. Essa formulação de inspiração aristotélica mostra que a verdade depende da estrutura

⁵⁴ Os três últimos parágrafos se baseiam em (FAGUNDES, 2015).

⁵⁵ Isso não significa que Tarski defenda uma teoria da verdade como correspondência. Haack (1976) escreve um artigo no qual chama a atenção para o fato de que a teoria de Tarski, embora compatível com uma teoria da verdade como correspondência, não se compromete com isso (O artigo é uma crítica a Popper que, segundo ela, entende a teoria da verdade de Tarski como correspondendista.). De acordo com ela, ao insistir na distinção entre metalinguagem e linguagem-objeto, o que o trabalho de Tarski pretende é oferecer uma condição material de adequação para qualquer teoria da verdade. Até mesmo uma concepção de verdade totalmente implausível como “‘p’ é verdadeira se, e somente se, o Papa afirma p” pode ser compatível com a condição proposta por Tarski. Além disso, o próprio Tarski não apresenta de nenhuma forma sua teoria como a defesa de uma correspondência entre frases e fatos.

semântica das frases. Sendo assim, a noção de verdade e a de significado andam juntas. Ora, compreender o significado de uma enunciação é compreender as suas condições de verdade.

Qualquer tentativa de definir a verdade, para o autor, deve ser abandonada, pois é sempre circular (SANCHES, 2002, p. 32), depende de uma concepção prévia de verdade. Essa é a razão pela qual a verdade deve ser tomada como conceitualmente primitiva. Davidson defende que, para interpretarmos qualquer frase em qualquer linguagem, precisaremos partir de um fundo de crenças majoritariamente verdadeiras. Caso contrário, a compreensão do significado de uma frase não é possível. A questão da verdade como conceito filosófico básico, portanto, é importante para que possamos compreender os procedimentos realizados pelo intérprete e que, em última análise, o levam a uma compreensão dos estados mentais do interlocutor.

Aqui, convém nos perguntarmos se é possível duas pessoas se comunicarem estando apoiadas sobre um conjunto de crenças compartilhadas e massivamente falsas. Isso não ocorre nas nossas vivências quotidianas, pois teríamos duas pessoas falando uma língua completamente incompreensível para o resto da humanidade. Podemos, contudo, tentar imaginar um cenário de ficção filosófica. As crenças deles são falsas porque são enganados da mesma maneira pelo mesmo Gênio Maligno. Nesse caso, ou nós estamos junto com eles no mesmo barco, e nada nos resta a não ser acreditar nas arbitrariedades do Gênio Maligno ou estamos do lado de fora, observando-os. Na segunda hipótese, o fato do discurso deles ser baseado em crenças majoritariamente falsas nos impedirá de compreendê-los, mas isso significa um completo descolamento entre o que dizem e o mundo, sua linguagem será algo completamente flutuante, nem mesmo eles terão um ponto a partir do qual construir seu discurso.

Recentemente, tivemos a notícia de duas Inteligências Artificiais que supostamente teriam desenvolvido uma linguagem própria e ininteligível e, em decorrência de seu comportamento assustador, teriam sido desligadas. É provável, contudo, que não tenha sido bem assim. Os robôs criaram uma linguagem própria para se comunicarem, mas os pesquisadores eram, sim, capazes de compreendê-la. Contudo, o objetivo da pesquisa era que os robôs desenvolvessem a habilidade de se comunicarem com humanos. Por isso, eles foram desligados e a pesquisa segue com novas inteligências artificiais ajustadas para os objetivos da pesquisa (GOMES, 2017).

A verdade, para Davidson, é uma correspondência entre o mundo e o que se diz, tem um aspecto objetivo, independente de nossas crenças. Contudo, essa correspondência não é uma relação direta entre proposições isoladas e o mundo. Dado que a compreensão do

significado se produz a partir da atribuição de crenças majoritariamente verdadeiras ao interlocutor e que de fato a comunicação existe, Davidson compreende que a verdade emerge a partir de um conjunto coerente de crenças. Para ele, “a coerência produz a correspondência” (DAVIDSON, 1987/2001b, p. 193). Com isso, a possibilidade de conhecimento se abre para os seres linguísticos.

O fato de que interpretamos um discurso garante que somos possuidores de um conjunto de crenças majoritariamente verdadeiras, embora muitas das nossas crenças também sejam falsas. Mas a verdade, para Davidson, não se caracteriza pela simples coerência. Ao contrário, é perfeitamente possível haver um conjunto de crenças coerentes entre si que sejam majoritariamente falsas. A verdade se caracteriza a partir da correspondência com a realidade. “Deveria estar claro que não espero definir a verdade em termos de coerência e crença. A verdade é lindamente transparente em comparação com a crença e a coerência, e a tomo como conceito primitivo” (DAVIDSON, 1987/2001b, p. 196, tradução nossa).

A importância da teoria coerentista da verdade no pensamento de Davidson é garantir que, embora um agente racional possa ter uma série de crenças falsas, e de fato as temos, ele não pode estar massivamente equivocado acerca da relação entre suas crenças e o mundo. Suas crenças não podem ser todas falsas nem podem ser majoritariamente falsas. A verdade de uma crença emerge de um conjunto coerente de crenças não porque a verdade seja simplesmente coerência, mas sim porque a única maneira de fundamentar a verdade de uma crença é pela sua participação em um conjunto de crenças verdadeiras. “O que distingue uma teoria coerentista é simplesmente a afirmação de que nada pode contar como razão para sustentar uma crença exceto outra crença” (DAVIDSON 1987/2001b, p. 198, tradução nossa).

Sanches (2002) caracteriza o pensamento de Davidson como *uma teoria da verdade como correspondência em conjunto com uma teoria da justificação como coerência*. A justificação tem um importante papel para que nossas crenças verdadeiras possam ser consideradas conhecimento⁵⁶. O que garante essa sustentação de umas crenças pelas outras é o fato de que os significados das palavras são aprendidos publicamente a partir das relações entre os falantes e os intérpretes e, para a interpretação ser possível, é necessário que o

⁵⁶ Há um longo debate acerca da noção de conhecimento como crença, verdadeira e justificada, e não entraremos aqui nesse debate. Porém, para termos uma noção do papel da justificação, suponhamos que alguém nos diga que não pegou o primeiro ônibus que passou para ir ao trabalho porque teve um misterioso pressentimento de que aquele ônibus seria assaltado e suponhamos que mais tarde, quando já estava no trabalho, ouviu a notícia de que aquele ônibus havia sido assaltado de fato. Você pergunta àquela pessoa como ela sabia que isso iria acontecer e ela responde que não sabia nada, apenas pressentiu. Nesse caso, embora possamos dizer que a pessoa teve uma crença verdadeira, tendemos a não chamar isso de conhecimento, pois a crença carece de uma justificação.

intérprete parta do pressuposto de um acordo fundamental entre suas próprias crenças e as crenças do falante acerca do que é verdadeiro.

Em outras palavras, se o intérprete atribui ao seu interlocutor um conjunto de crenças majoritariamente falsas, então suas frases podem se referir a tudo, exceto ao mundo que os cerca. Dessa forma, não é possível atribuir nenhum significado ao que o falante diz. Ele pode estar falando de qualquer coisa, não há parâmetro nenhum ao intérprete e a interpretação não é possível. De fato, afirma Davidson: “o acordo disseminado é o único cenário [“background”] possível no qual as disputas e erros podem ser interpretados” (DAVIDSON, 1975/1984d, p. 153, tradução nossa).

Ao longo do tempo, no trabalho de Davidson, a rede de relações complexas entre os pensamentos de um sujeito parece ganhar um papel ainda mais relevante em relação à questão da interpretação. Em “Interpretation: hard in theory, easy in practice” (1999), Davidson se depara com a questão acerca de algo que considera tão cotidiano e natural como a interpretação ser também algo tão difícil de descrever a partir de uma perspectiva teórica, mas enfatiza a importância das relações que os estados mentais do agente mantêm entre si.

[...] seria tolice subestimar o quanto é difícil descrever como nós detectamos os motivos e pensamentos dos outros e compreender o que dizem. Uma coisa que torna difícil conceber uma abordagem teórica da interpretação é a complexidade das interdependências entre as várias atitudes e a extensão na qual o conteúdo de um só pensamento ou expressão descansa em seu lugar em uma rede de mais pensamentos e expressões. Essas interdependências implicam que a compreensão de qualquer crença, intenção, desejo, ação ou proferimento de um agente é sempre contingente ao conhecimento de uma quantidade verdadeiramente vasta do resto dessas atitudes do agente. (DAVIDSON, 1999, p. 31, tradução nossa)

2.1.3. Tradução e interpretação

A ideia do intérprete radical proposta por Davidson é de que, para que seja possível atribuir sentido ao discurso de alguém, é necessário que se atribuam também atitudes proposicionais a essa pessoa. A tese de Davidson, portanto, é que a fonte do significado é a interpretação bem sucedida. Davidson (1986) cita os “malapropismos”, isto é, usos impróprios de expressões parônimas. Por exemplo, às vezes, os professores de filosofia encontram no trabalho de algum estudante de graduação a expressão “preposição” (palavra que estabelece uma ligação entre dois elementos de uma frase) no lugar de “proposição” (o conteúdo de uma frase, ao qual é possível atribuir um valor de verdade). Na primeira vez que isso ocorre, o professor pode ficar um pouco confuso, mas logo passará a ter uma compreensão automática desse erro, apenas fazendo a correção para que o estudante não volte

a trocar as expressões. Para isso, o professor atribui ao estudante algumas atitudes proposicionais, inclusive o desejo de obter nota suficiente para aprovação. Assim, ainda que o estudante troque as palavras, o professor é capaz de compreender sem dificuldade o que se quer dizer, o que revela uma interpretação bem sucedida daquela situação, apenas corrigirá o estudante para evitar a repetição do malapropismo.

Conforme Davidson, em uma situação de interação linguística, os interagentes partem de uma teoria anterior acerca das intenções uns dos outros, ou seja, acerca do significado que querem dar a seus termos. Ambos querem se comunicar, querem ver um sentido nas palavras dos outros e querem dar um sentido às suas próprias palavras perante os outros. Porém, nas situações reais de comunicação, essa teoria é substituída por uma teoria transitória [“*passing theorie*”], em que a teoria anterior é ajustada e os acordos linguísticos se formam. Isso se torna evidente nos casos de metáforas e malapropismos, pois são situações nas quais a teoria anterior é quebrada, ou seja, o intérprete possuía certas expectativas em relação ao falante, mas essas expectativas são quebradas.

O intérprete vem para a ocasião da fala armado com uma teoria que diz a ele (ou assim ele acredita) o que uma fala arbitrária do falante significa. O falante então diz algo com a intenção de que isso vá ser interpretado de uma certa forma, e a expectativa de que assim será interpretado. Na realidade, essa forma não é dada pelo intérprete da teoria. Mas o falante é, no entanto, compreendido; o intérprete ajusta sua teoria de modo que produza a interpretação pretendida pelo falante (DAVIDSON, 1986, p. 259, tradução nossa).

A teoria transitória aproxima as intenções do falante e do intérprete ao longo da situação real de comunicação (SILVA FILHO, 2001). Eles não possuem uma convenção ou uma teoria completamente compartilhada antes da comunicação. Porém, no decorrer do processo, seus teorias semânticas se tornam cada vez mais próximas. O que eles compartilham é a teoria transitória, e é essa teoria que permite a compreensão da linguagem dentro do contexto da comunicação. A teoria transitória é necessária, segundo Davidson, para que seja possível interpretar uma enunciação em uma situação específica, pois as falas não possuem um sentido solto, isolado das situações concretas de comunicação.

Uma vez que o discurso da pessoa tem um sentido, duas conclusões podem ser extraídas: primeiro, que as atitudes proposicionais atribuídas pertencem de fato ao interlocutor; segundo, que as crenças do interlocutor são majoritariamente verdadeiras. Essa necessidade que o intérprete tem de considerar como verdadeiras a maior parte das crenças do falante é denominada “princípio de caridade”.

Ao mesmo tempo, um requisito importante para que a interpretação seja possível, conforme Caro (1999), é que haja uma compatibilidade entre o externismo de Davidson e a

autoridade de primeira pessoa acerca dos estados mentais⁵⁷. Só é possível interpretar o discurso de alguém se pressupomos que essa pessoa sabe o que significam as suas próprias palavras, que ela crê que esses significados são associados por todos às palavras que ela usa e que ela quer dizer aquilo que efetivamente diz⁵⁸.

Há um aspecto privado na crença, que é simplesmente o fato de ela ser particular, de ela ser a crença de um sujeito. Mesmo assim, isso não faz com que ela seja acessível apenas ao sujeito. “As atribuições de crença são tão publicamente verificáveis quanto as interpretações, baseado nas mesmas evidências: se nós podemos entender o que uma pessoa diz, nós podemos saber no que ela acredita” (DAVIDSON, 1975/1984d, p.153, tradução nossa).

Davidson (1994a, p. 125) considera que as linguagens naturais são acessíveis ao conhecimento humano. Ainda que haja uma série de enganos na comunicação, a quantidade de acertos e concordâncias supera esses erros, tornando possível alguma compreensão e o pensamento sobre a realidade que nos cerca. O discurso de um estrangeiro proferido em uma língua desconhecida, como vimos, pode ser compreendido pelo intérprete radical. O processo de compreensão se caracteriza pela construção gradual de conexões entre o discurso e o mundo. Por isso, a referência dos termos singulares e dos predicados deve ser aprendida por meio das relações que o falante e o intérprete possuem no contexto da interação social.

Como são feitas essas conexões entre a linguagem e o mundo? Elas devem ser feitas a partir das evidências observáveis, tudo o que esteja publicamente acessível ao intérprete: quem é o falante, o contexto, sons que são repetidos em determinadas situações, tentativas, por parte do intérprete, de imitar aqueles sons e relacionar às situações do mundo, observações detalhadas do comportamento do falante. Assim, o método da tradução radical proposto por Quine valerá para a interpretação radical, com algumas diferenças.

Como a noção davidsoniana de interpretação radical se compara à tradução radical de Quine? O intérprete necessita formular uma teoria acerca do que o seu interlocutor está dizendo. Para isso, buscará relacionar os sons proferidos e o comportamento do falante ao mundo. Primeiro, o intérprete precisará identificar as frases proferidas em forma de sons. Em seguida, precisará identificar, por meio do comportamento do agente, a rejeição a

⁵⁷ Ver: seção 2.2.

⁵⁸ Vale a pena, aqui, mencionar essa passagem de Tomasello (2014, p.1, tradução nossa): “O pensamento poderia parecer uma atividade completamente solitária. Assim é para as outras espécies animais. Mas para humanos, o pensamento é como um músico de jazz improvisando um novo refrão na privacidade de seu quarto. Ela é solitária, tudo bem, mas em um instrumento feito por outros para esse propósito geral, após anos tocando com e aprendendo com outros participantes, em um gênero musical com uma rica história de refrãos lendários para uma audiência imaginária de aficionados em jazz.”

determinadas frases, as quais serão tomadas como falsas e a aceitação de outras, as quais serão tomadas como verdadeiras. Depois, precisará relacionar essas frases aceitas ou recusadas às situações reais do mundo e, dessa forma, poderá construir ligações entre os sons proferidos e o mundo que os cerca (ao intérprete e ao falante).

O intérprete, ao notar que o agente regularmente aceita ou rejeita a frase “o café está pronto”, quando o café está ou não está pronto (contudo, tentativamente à espera de resultados relacionados), se esforçará para conseguir uma teoria da verdade que diga que um enunciado [“utterance”], feito pelo agente da frase “o café está pronto” é verdadeiro se e somente se o agente puder observar que o café está pronto no momento do enunciado⁵⁹. (DAVIDSON, 1990/2002a, p. 100, aspas internas do original)

O método, portanto, é bastante similar ao da tradução radical. Porém, Davidson coloca sua ênfase na relação do falante e do intérprete com o mundo circundante, e não com os estímulos sensoriais. Isso é importante para sua concepção externista. O que é relevante para explicar a interpretação não são os estímulos sensoriais na mente do intérprete, mas sim o mundo externo, aquilo que circunda o contexto da comunicação. Quine estava interessado em desenvolver uma concepção empirista, mas até o empirismo é rejeitado por Davidson por ser também uma concepção subjetivista e, no fim das contas, um tipo de defesa do Teatro Cartesiano. Há, portanto, uma diferença de ênfase que se estabelece a partir do empirismo de Quine e da negação por parte de Davidson desse empirismo.

Para diferenciar a interpretação radical da tradução radical, Davidson (1973/1984b, p. 129) explica que uma teoria da interpretação envolve duas linguagens, uma vez que a interpretação é feita numa linguagem conhecida pelo intérprete. Uma teoria da tradução, por sua vez, envolve três linguagens: a linguagem a ser traduzida, a linguagem na qual a tradução é dada e a linguagem da teoria, que diz quais expressões da linguagem na qual é dada a tradução traduzem quais expressões da linguagem a ser traduzida. Ou seja, embora normalmente a linguagem da teoria corresponda à linguagem na qual é dada a tradução, também poderia não corresponder. Por isso, é possível haver a tradução sem que haja a interpretação, ou seja, também é possível a tradução ser feita sem que haja compreensão do significado das palavras. De todo modo, nessas circunstâncias, qualquer pessoa que compreenda a linguagem da teoria estará apta a compreender também a linguagem a ser traduzida, mas é necessário haver essa convergência entre os conhecimentos prévios do tradutor e a linguagem da teoria, transformando o processo de tradução em interpretação.

⁵⁹ O termo “enunciação” seria, aqui, preferível a “enunciado”, pois estamos diante de uma ação. Porém, mantivemos o termo presente na tradução utilizada.

Tornar possível a interpretação pressupõe algumas similaridades entre a estrutura semântica da língua do falante e do intérprete. Por exemplo, “a interpretação dos proferimentos das frases complexas dependerá sistematicamente da interpretação dos proferimentos de frases mais simples” (1973/1984b, p.130, tradução nossa). Davidson considera também que será necessário ao intérprete pressupor que o falante estrangeiro possui o mesmo uso das constantes lógicas que ele possui. Isso permitirá a identificação dos termos dentro das frases (pressupõe-se a existência de termos singulares e predicados na linguagem alheia, por exemplo) e a maneira como essas frases se concatenam (pressupomos que há conectivos lógicos passíveis da mesma interpretação verofuncional).

A lógica de primeira ordem é tomada como uma espécie de suporte, não em busca de uma relação unívoca entre as constantes lógicas e o discurso do falante, mas como um pressuposto racional para a interpretação. Considerar-se-á, por exemplo, que, se o falante sustenta como verdadeira a frase “ p e q ”, não rejeitará a frase “ p ”. Isso permitirá a identificação de quantificadores, termos singulares, predicados, conectivos... (DAVIDSON, 1973/1984b, p. 136). O intérprete também deverá se concentrar nas frases que possuem indexicais, para identificá-los, trazendo limites semânticos para os termos. Por último, precisará lidar com as frases acerca das quais parece não haver um acordo, buscando cuidadosamente o que significam. O método pressupõe a existência do fundo de acordos massivos entre intérprete e falante. Essa pressuposição do “princípio de caridade” é necessária até mesmo para que consideremos a fala do interlocutor como sendo significativa, caso contrário, não há razão para o esforço em se interpretar. Sem isso, o esforço pela interpretação perde completamente o sentido.

2.2. Princípio de caridade e o terceiro dogma do empirismo

O princípio de caridade é um requisito para a compreensão de qualquer comportamento – inclusive o comportamento verbal - que consiste em maximizar a racionalidade do interlocutor e interpretar os possíveis erros e desacordos minimizando a chance de equívoco alheio. O princípio postula que o intérprete considere o falante como alguém racional, comprometido com dizer coisas nas quais de fato acredita, além de verdadeiro em suas palavras. Se partisse do pressuposto de que o outro está proferindo um

discurso falso, o intérprete se tornaria impermeável àquilo que está sendo dito, fechando-se dentro de sua própria ignorância⁶⁰.

Supondo, por exemplo, um leitor que pretende entender o *De Anima*, de Aristóteles. Se, diante da primeira dificuldade, o leitor concluir que Aristóteles não sabe se expressar e que tudo o que diz são tolices, certamente não poderá prosseguir a sua leitura e jamais encontrará qualquer sentido nas palavras ali escritas. Por sua arrogância, estará impedido de experimentar qualquer aventura junto a um texto fundamental da história do pensamento ocidental. Apenas adotando o princípio de que Aristóteles foi racional na maior medida possível e que suas frases são majoritariamente verdadeiras é que o texto poderá ser compreendido. Esse exemplo pode ser ampliado para qualquer texto ou qualquer discurso, seja ele filosófico ou não, escrito ou falado, em verso ou em prosa.

O princípio de caridade é definido por Caro (1999, p.16, tradução nossa) no pensamento de Davidson como “a necessária atribuição ao falante de certo grau de consistência lógica junto com grande quantidade de crenças verdadeiras”. A caridade, na concepção de Davidson, “não é uma opção” (1974/1984c, p.197, tradução nossa), mas uma condição de possibilidade para a interpretação (outra condição seria uma teoria da verdade). “A caridade é forçada sobre nós, quer gostemos disso ou não, se queremos entender os outros, precisamos tomá-los como certos na maioria das questões.” (1974/1984c, p. 197, tradução nossa)⁶¹.

No artigo mencionado acima, Davidson está preocupado em buscar uma superação para o dualismo esquema/conteúdo - terceiro e último dogma do empirismo em sua concepção (depois dos dois dogmas de Quine) e que, uma vez superado, levará ao fim do empirismo. O autor considera que o princípio de caridade tem um papel importante nessa superação.

Dualismo esquema/conteúdo é uma ideia que possui raízes no pensamento kantiano. De um lado, estariam os esquemas conceituais, de outro, as experiências sensíveis. Os esquemas conceituais seriam categorias capazes de classificar, organizar e dar sentido às experiências, adaptando as suas anomalias e modelando a apreensão da realidade, tornando-a algo relativo a eles. Por isso, Davidson afirma que aceitar esse tipo de dualismo é aceitar a possibilidade de existência de diferentes esquemas conceituais (relativismo de esquemas

⁶⁰ O arrogante, por sua ignorância, não é capaz de alcançar qualquer compreensão do outro. Como se lê no “Banquete”, de Platão: “Os ignorantes [...] não filosofam e não desejam se tornar sábios. É isso justamente que é deplorável na ignorância: não se é belo, nem bom, nem inteligente e, no entanto, se acredita sê-lo. Não se deseja uma coisa quando não se sente sua falta.” (*Apud* DROZ, 1992/1997, p. 42)

⁶¹ Aqui, a caridade se apresenta como um princípio instrumental. Ela é necessária para que um discurso possa ser compreendido. Podemos nos perguntar, contudo, se as palavras do interlocutor são realmente majoritariamente verdadeiras.

conceituais). Para Davidson, mente, linguagem e mundo são compreendidos de maneira inter-relacionada e pressupor um dualismo entre esquema/conteúdo acaba por produzir uma separação a partir da qual a mente não pode ser compreendida. A mente, na concepção dele, é constituída de linguagem e a linguagem é constituída na relação com o mundo.

Para Kuhn, por exemplo, os períodos em que uma comunidade científica trabalha dentro de um determinado esquema conceitual (um paradigma⁶², na terminologia kuhniana) é completamente incomparável a um período científico no qual o esquema conceitual é outro, como se, nos diferentes períodos, a própria realidade percebida fosse diferente. Dentro dessa concepção, os cientistas estariam tão submersos em seus esquemas conceituais que seria impossível uma comparação entre os diferentes momentos da história da ciência, pois isso exigiria a adoção de um ponto de vista neutro, externo a qualquer esquema conceitual, o que Kuhn recusa. Sequer seria possível uma tradução a partir de um esquema conceitual ao outro, pois até os significados dos termos seriam diferentes.

Davidson, por outro lado, entende não ser possível haver diferenças profundas entre modos de olhar a realidade que justifiquem o relativismo de esquemas conceituais. Por mais que existam desacordos entre as pessoas, entre grupos de pessoas e entre as diferentes culturas, esses desacordos se desenrolam sobre o pano de fundo de uma grande quantidade de crenças comuns⁶³. Seu argumento contra o dualismo de esquemas conceituais requer que entre duas línguas, por mais que elas estejam distantes uma da outra, a interpretação deva sempre ser, de algum modo, possível. Se não fosse possível realizar uma tradução, ainda que mais ou menos imprecisa, entre duas línguas, pensa Davidson, nesse caso estaríamos diante de dois esquemas conceituais diferentes.

De acordo com McDowell (2005), essa superação do dualismo entre esquema conceitual e conteúdo é importante no pensamento de Davidson porque tal dualismo está

⁶² Os paradigmas de Kuhn (1962/1998, p.30) são conjuntos de teorias, leis, métodos, um vocabulário, instrumentos e princípios compartilhados que a comunidade científica envolvida com uma disciplina adota durante certo período histórico e que permite o desenvolvimento da pesquisa científica dentro daquela disciplina. Para Kuhn, a ciência não progride fora de um paradigma, pois é preciso haver uma base de acordo sobre a qual a ciência possa crescer. Os paradigmas mudam ao longo da história da ciência e não é possível compará-los. Os termos adquirem mudanças de uso nos novos paradigmas, ganhando significados diferentes. “O resultado inevitável é o que devemos chamar, embora o termo não seja bem preciso, de um mal-entendido entre as duas escolas competidoras” (1962/1998, p. 189). Assim, poderíamos compreender a incomensurabilidade como intradutibilidade. Hoje, na filosofia da ciência, há certo consenso de que houve um exagero relativo à incomensurabilidade. Conseguimos compreender as teorias científicas da antiguidade, por exemplo.

⁶³ Podemos derivar disso a tese de que toda a humanidade possui um fundo de crenças compartilhadas. Isso provavelmente está relacionado à forma como evoluímos. De acordo com Caro (1999, p. 16, tradução nossa), Davidson afirma que o processo de triangulação – envolvendo falante, intérprete e mundo – só é possível “porque a evolução e a aprendizagem fizeram muito similares os modos nos quais os diferentes seres humanos classificam os estímulos”. De fato, o desenvolvimento da linguagem, na concepção de Davidson, tem relação com as necessidades adaptativas de nossos ancestrais. Isto será desenvolvido na última parte deste capítulo.

firmemente vinculado a uma concepção da mente e de seu lugar na natureza que é construída desde a Modernidade e que Davidson pretende desconstruir. Em outras palavras, ao desfazer esse dualismo, a mente perde seu papel de organizadora externa das experiências sensíveis, podendo ocupar um novo lugar junto a elas⁶⁴.

Vale notar que, ao extinguir o dualismo entre esquema/conteúdo, Davidson tem um ganho também em termos de simplicidade epistemológica e ontológica, uma vez que a relação entre mundo e pensamento passa a ser direta, sem intermediários. Não há, na concepção de Davidson, objetos perante a mente. O objeto do pensamento é o próprio objeto do mundo (SMITH, 2005). A crença possui uma relação causal direta com outras crenças e com o mundo, o que traz um suporte empírico aos pensamentos acerca do mundo.

Vimos que a concepção de verdade de Davidson é fortemente inspirada na Convenção-V de Tarski, que envolve a construção em metalinguagem de frases que traduzem as frases da linguagem objeto. As noções de verdade e tradução, portanto, andam de mãos dadas no sistema davidsoniano. Porém, argumenta Davidson (1974/1984c), se houvesse esquemas conceituais diferentes, as noções de tradução e de verdade estariam dissociadas, não haveria como realizar uma tradução entre duas línguas que estivessem vinculadas a diferentes esquemas conceituais. Davidson afirma que, se houvesse uma língua completamente intraduzível não poderia ser reconhecida como uma língua⁶⁵. Por outro lado, se a língua é parcialmente intraduzível, isso pode ser ajustado sobre um pano de fundo de acordos massivos entre os falantes das diferentes línguas, conforme o princípio de caridade. Ou seja, no lugar de simplesmente rotular como falsa ou absurda a fala do outro, buscamos a caridade, sem a qual não existiria diálogo.

Como mencionado, seguindo o método da tradução radical de Quine, Davidson defende que a interpretação radical requer que se atribua ao interlocutor a lógica de predicados de primeira ordem com identidade. Isso, para ele, consiste em atribuição de racionalidade a qualquer ser capaz de comunicação linguística. Assim, ao interpretar um discurso, atribuir-se-á ao conjunto de suas crenças uma consistência, considerando não haver grandes contradições entre elas. Quine (1960/2010, p.88) está convencido de que só é possível fazer uma tradução se atribuirmos ao estrangeiro as leis lógicas que atribuimos a nós mesmos. É preciso pressupor, por exemplo, que há um “sim” e um “não” e que há uma contradição

⁶⁴ Convém observar que, nesse artigo, McDowell argumenta que a superação do dualismo entre esquema conceitual e conteúdo sensível não é suficiente para a transformação desejada por Davidson no pensamento filosófico.

⁶⁵ Se houvesse esquemas conceituais diferentes, isso significa que haveria uma linguagem intraduzível. Embora o argumento seja hipotético, ele parece ter uma conclusão oposta ao que de fato conclui. Desenvolveremos um pouco isso na próxima seção.

entre eles; também que as constantes lógicas passam no teste da interpretação radical. Se há uma aparente violação da lógica de predicados de primeira ordem, na concepção de Quine, isso provavelmente significa que há problemas no trabalho do tradutor, necessitando ser refeito.

Um exemplo interessante dado por ele é o da língua espanhola com seu “No hay nada”. Em língua portuguesa, temos também esse tipo de construção como em “não veio ninguém”. Conforme Quine (1960/2010, p. 89): “Amantes do paradoxo podem representá-lo desacatando a lei da dupla negação. Tradutores soberbos podem supor ‘não’ e ‘nada’, nesse contexto, como metade de um negativo” (aspas duplas da tradução substituídas por aspas simples). Mas isso seria uma solução precipitada. Diante de uma situação aparentemente contraditória, o tradutor deverá ajustar seu trabalho, sempre levando em consideração que a língua a ser traduzida se constrói a partir dos mesmos princípios lógicos. É muito mais provável, segundo ele, que a aparente falsidade no discurso do estrangeiro seja sinal de erro de tradução do que de tolice do estrangeiro. Para que seja possível traduzir um discurso estrangeiro, nessa concepção, uma alta dose de caridade é o remédio.

Quine (1960/2010, p. 88) considera fundamental o combate à ideia de que a fala de uma pessoa não tenha como fundamento a lógica, ideia chamada por ele de “mito da mentalidade pré-lógica”. Seria absurdo, para ele, considerar que os falantes de certa língua desconhecida defendessem simultaneamente a proposição p e a proposição $\text{não-}p$, pois isso iria contra nossas intuições mais básicas. Esses mesmos critérios apresentados por Quine para a tradução, segundo Davidson, devem ser seguidos para que a interpretação seja possível, isso faz parte da adoção do princípio de caridade.

Por outro lado, Davidson deixa claro que tem uma posição distinta da de Quine em relação aos esquemas conceituais, pois vê em Quine uma associação entre esquemas conceituais e linguagens, sendo que os primeiros estabelecem uma relação com a experiência, moldando-a. A posição de Quine acerca da relação entre linguagem e mundo é de “relativismo ontológico”, pois para ele há uma dificuldade, no processo de tradução, de estabelecer com certeza a referência do que diz o falante de uma língua estrangeira (LOUX; SOLOMON, 2006). Para Davidson, por outro lado, há uma relação direta entre experiência e mundo. Há também uma relação entre experiência e crença que dá suporte à verdade das crenças.

2.2.1. *Relativismo ou realismo?*

Malpas (2005) defende que Davidson está comprometido com uma posição denominada “realismo cotidiano”. Dentro do debate filosófico tradicional entre realismo e idealismo, “O realismo sustenta que o mundo físico existe independentemente do pensamento e da percepção humana. O idealismo nega-o — afirma que o mundo físico é de algum modo dependente da atividade consciente dos seres humanos” (OKASHA, 2017, tradução nossa). Porém, o realismo de Davidson não cabe nesse debate, pois parte de outro ponto. O debate tradicional provém de uma ruptura entre sujeito e objeto que é justamente o que Davidson rechaça. Davidson pensa a sua filosofia a partir do “envolvimento ordinário, cotidiano com o mundo” (MALPAS, 2005, p. 51), o que situa seu realismo fora da discussão metafísica predominante. O realismo cotidiano, portanto, consiste na defesa da verdade de nossas crenças a partir do fato de que elas se constroem na relação com o mundo. Nem o mundo físico independe do pensamento nem é constituído por ele. O pensamento, sim, é constituído pelo mundo físico e é parte dele⁶⁶.

Esse ponto pode ser uma chave para se compreender o mental e sua inserção no mundo material. Há alguns sentidos nos quais se pode dizer que o pensamento é parte do mundo físico. Por exemplo, há o sentido evolutivo, pois o pensamento surge como parte de um processo natural e é algo organicamente muito dispendioso para ter se desenvolvido sem que tenha aumentado a aptidão da espécie. Ora, o que se desenvolve por um processo evolutivo está submetido ao crivo da seleção ambiental, não está encarcerado no sujeito. Contudo, o sentido mais importante em que isso pode ser dito dentro da argumentação davidsoniana é o seguinte: o pensamento é linguístico, depende da relação direta do sujeito com a realidade física e social.

Diante do princípio de caridade de Davidson, uma questão é saber se ele é apenas um método instrumental para que possamos dar sentido ao discurso do outro ou se de fato cada ser linguístico realmente possui todas essas crenças que lhe são atribuídas. Uma coisa é reconhecermos que precisarmos atribuir alto grau de racionalidade e de verdade às crenças dos outros para lhes podermos compreender. Outra coisa é defendermos que os seres humanos de fato são altamente racionais e estão massivamente corretos.

Acerca desse tema, Tennant (1999) apresenta a seguinte citação de Davidson:

⁶⁶ Aqui, não estamos falando de pensamento no sentido de Frege (2002) de algo pertencente a um terceiro domínio, ao mesmo tempo objetivo e não apreensível pelos sentidos.

O processo de elaboração de uma teoria da verdade para uma língua nativa poderia grosso modo ser esboçado como se segue. Primeiro, nós procuramos a melhor forma de adaptar a nossa lógica até onde requerido para ter uma teoria que satisfaça a convenção V na nova linguagem; isso pode significar ler a estrutura lógica da teoria de quantificação de primeira ordem (mais identidade) na linguagem, não pegando as constantes lógicas uma a uma, mas tratando essa parte da lógica como uma rede a ser adaptada à realidade de uma só vez (DAVIDSON *apud* TENNANT, 1999, p. 71, tradução nossa).

Tennant, então, propõe que a exegese correta desse trecho inclui a compreensão que a atribuição da lógica de predicados é feita *somente* até onde necessário para que a convenção-V seja satisfeita, e isso é algo bem fraco. Assume-se que há operadores lógicos análogos aos nossos e estruturas sintáticas que fazem composições das expressões da linguagem a ser interpretada. Porém, isso não significa que seja necessário que os falantes da língua a ser interpretada concordem com todas as regras da nossa lógica clássica. Para Tennant (1999, p. 74), a proposta de Davidson é de um método minimalista. Achar que o falante a ser interpretado compartilha conosco de toda a lógica de predicados com identidade formalizada é assumir uma estrutura de verdades completamente independentes da mente, mas que podemos descobrir. Tennant não acha que essa seja a ideia que Davidson pretende defender.

Por outro lado, mesmo concordando com a interpretação radical, Tennant defende os esquemas conceituais e a possibilidade de tradução entre eles. Para ele, a interpretação radical não tem um papel relevante na superação do dualismo esquema/conteúdo. O autor oferece um experimento de pensamento em que pede para imaginarmos extraterrestres com sistemas sensoriais completamente diferentes dos nossos⁶⁷. Precisaremos imaginar também que nós e eles desenvolvemos instrumentos tecnológicos que nos permitiriam preencher as lacunas das diferenças sensoriais entre nós, transformando e adaptando os diferentes modos de enxergar a realidade. Além disso, nós e eles teríamos criado extensões científicas de nossos aparatos linguísticos tornando possível algum tipo de tradução da língua extraterrestre para a humana. Nesse caso, conforme Tennant, teríamos algum nível de tradutibilidade entre diferentes esquemas conceituais. Com esse argumento, ele pretende defender um relativismo de esquemas conceituais.

O autor apresenta, ainda, outro argumento muito simples: “poderia haver esquemas conceituais alternativos precisamente porque eles são inacessíveis” (TENNANT, 1999, p. 78, tradução nossa). Em outras palavras, se Davidson está certo ao argumentar que não podemos

⁶⁷ Com *qualia* distintos, mas a concepção de *qualia* de Tennant não parece ser de estados causalmente isolados dos outros estados mentais nem da linguagem.

atribuir linguagem a uma entidade se há intraduzibilidade completa, isso não significa que tal entidade não possua linguagem. Aquele ser poderia ter um esquema conceitual distinto, mas ser racional e se comunicar linguisticamente, ainda que nós não fôssemos capazes de detectar isso. O que impediria uma linguagem completamente intraduzível de ser uma linguagem?

Qual o reflexo disso para o princípio de caridade? O princípio de caridade parece pressupor o mais alto grau de correlação entre linguagem, falante, intérprete e mundo. Ao adotar esse princípio, uma das coisas que o intérprete faz é pressupor que o falante relaciona determinados termos de sua linguagem a determinados termos do mundo da mesma forma que ele, intérprete. Esse princípio, portanto, se aplica a seres que compartilham de uma série de crenças (ou que compartilham de um esquema conceitual, se aceitamos que existem esquemas conceituais). Se houvesse alienígenas com o aparato sensorial completamente diferente do nosso, então, faria algum sentido aplicar a eles o princípio de caridade?

No modo de entender de Tennant, é possível aceitarmos a possibilidade de tradução mútua entre todas as linguagens, mesmo que alienígenas, e aceitar o princípio de caridade, mas mantendo uma abordagem relativista na relação entre esquemas conceituais e realidade, admitindo a existência de esquemas conceituais mediando a relação entre sujeito e mundo. Dessa forma, o autor pretende evitar o antropocentrismo acerca da linguagem e do pensamento proposicional.

Todas as linguagens humanas incorporam o mesmo esquema conceitual. Não se segue, contudo, que todo código possível para a comunicação de pensamento representativo do mundo externo tenha que estar na mesma classe de equivalência das linguagens usadas por seres humanos (TENNANT, 1999, p. 89, tradução nossa).

Como denunciado por Davidson, a própria noção de esquema conceitual conduz ao relativismo. Conforme Baghramian e Carter (2016, tradução nossa):

Relativismo, grosseiramente colocado, é a visão de que verdade e falsidade, certo e errado, padrões de raciocínio e procedimentos de justificação são produtos de diferentes convenções e estruturas [“frameworks”] de avaliação, e sua autoridade é confinada ao contexto que lhes dá origem. Mais precisamente, “relativismo” circunscreve visões que mantêm que – em um alto nível de abstração – ao menos uma classe de coisas tem as propriedades que tem (p.ex., bonito, moralmente bom, epistemicamente justificado) não *simpliciter*, mas apenas relativamente a uma dada estrutura de avaliação (p.ex., normas culturais locais, padrões individuais), e correspondentemente, que a verdade de alegações que atribuem essas propriedades só se mantêm se a estrutura de avaliação relevante for especificada ou fornecida. Relativistas caracteristicamente insistem, além disso, que se algo é apenas relativo, então não pode haver ponto de vista independente da estrutura a partir do qual se possa estabelecer se a coisa é realmente aquilo.

O relativismo condiciona a verdade ao esquema conceitual, o que a afasta do contato com a realidade. Será que poderíamos ao mesmo tempo admitir a possibilidade de que os extraterrestres com aparatos sensoriais distintos dos nossos desenvolvam uma linguagem proposicional e recusar o relativismo de esquemas conceituais?

Haack (1996) apresenta o relativismo como uma família de teses segundo a qual “algo é relativo a algo”. Para ela, contudo, há aspectos do mundo que independem de nós e de como nós pensamos que ele seja (embora o mundo não seja uma totalidade fixa independente do mental, ou seja, ela defende a realidade do mental). Por exemplo, uma pedra é algo que está no mundo independentemente do que quer que pensemos acerca dela. Ela denomina sua posição de realismo inocente [“innocent realism”]. Sua posição, grosso modo, é de que há muitos vocabulários e descrições diferentes que podem dizer o mesmo sobre o mundo, sendo compatíveis e podendo ser verdadeiros simultaneamente. Por outro lado, descrições contraditórias da realidade não poderiam ser simultaneamente compatíveis.

Essa perspectiva, ainda que “inocente” pode ser esclarecedora para compreendermos o tipo de realismo de Davidson (realismo quotidiano, na denominação de Malpas). Como seria possível descrições diferentes da mesma realidade serem ambas verdadeiras? Pelo fato de que elas foram construídas a partir da relação com a realidade. Sem essa relação, não haveria linguagem alguma para descrever coisa nenhuma. Então, por mais diferentes que duas linguagens sejam entre si, ambas conseguirão estabelecer pontos de contato com a realidade. Haack enfatiza que essa proposta de realismo não se compromete com uma perspectiva da verdade como correspondência sendo, portanto, compatível com o coerentismo de Davidson.

A evolução deu ao ser humano uma série de instrumentos comuns para lidarmos com o mundo, instrumentos que nos permitiram a construção da linguagem e de outros instrumentos externos a nós para olhar e decifrar essa realidade, mas sobre uma base comum, mais ou menos compartilhada por todos. Se, na linha de Davidson, recusarmos o dualismo esquema/conteúdo, caímos em um antropocentrismo? Essa parece ser, então, uma motivação de Tennant para defender a retomada desse dualismo. Porém, convém investigar a possibilidade de recusar o dualismo esquema/conteúdo sem precisar recusar junto a possibilidade de que extraterrestres com aparatos cognitivos completamente diferentes dos nossos possam desenvolver uma linguagem, ou seja, sem precisar cair em uma visão completamente antropocêntrica.

Se houvesse extraterrestres com aparatos sensoriais completamente diferentes dos nossos, em princípio, nada impediria que eles tivessem linguagem, ainda que não tivéssemos recursos para realizar a tradução. Mas como a linguagem teria se desenvolvido para eles? Ora,

se aceitamos a argumentação de Davidson, podemos considerar que a linguagem para eles se desenvolveu da mesma maneira que para nós, ou seja, a partir da relação entre falantes e mundo, por meio de relações causais que envolvem a própria realidade. O próprio argumento de Tennant parece nos conduzir a essa conclusão, já que, no experimento de pensamento proposto por ele, o desenvolvimento dos instrumentos científicos acaba por permitir algum nível de tradução entre nossa língua e a língua dos extraterrestres. Nesse sentido, embora nós e eles tenhamos aparatos cognitivos diferentes, há a possibilidade de nos ajustarmos justamente porque nosso conhecimento e o deles foi construído sobre as mesmas bases: as relações com o mundo real (certamente, o fato de termos evoluído em planetas diferentes irá dificultar bastante as coisas, mas a existência de um mundo compartilhado é pré-requisito para a possibilidade de tradução). Por isso, nós e eles também compartilhamos de um fundo de crenças a partir do qual, com uma série de ajustes, se torna possível a compreensão mútua.

Para compreender isso, não precisamos postular diferentes esquemas conceituais. O aparato sensorial deles pode ser diferente do nosso, mas a linguagem deles foi constituída por meio do contato direto com a realidade, mesmo que essa realidade seja muito diferente da nossa, seja lá que instrumentos eles possuem para fazer esse contato. Por mais que o contato deles com o mundo possa se dar por outras vias, isso não significa que eles constituirão uma perspectiva da realidade completamente diferente da nossa e que seus conceitos serão completamente distintos. Tanto é que, com o desenvolvimento científico e tecnológico, no experimento de pensamento de Tennant, a tradução começa a se tornar possível. Assim, o experimento de pensamento não conduz ao relativismo de esquemas conceituais e, se for unido ao pensamento de Davidson, nos conduz a um tipo de realismo: esse realismo davidsoniano que compreende as crenças como sendo construídas a partir do envolvimento direto com o mundo.

O princípio de caridade, portanto, pode ser conciliado com a ideia de que outros seres, constituídos pela sua relação com o mundo de forma completamente diferente da nossa, possam também ter desenvolvido uma linguagem, ainda que não sejamos capazes de reconhecer essa linguagem. Essa ideia, Tennant também pretende defender, mas à custa da adoção de um relativismo de esquemas conceituais. Mas isso não é necessário. Podemos aceitar essa ideia e, ao mesmo tempo, manter a rejeição do que Davidson considera o terceiro dogma do empirismo, adotando um tipo de realismo bastante fundamentado nas nossas experiências comuns.

2.3. Triangulação

Marcia Cavell (2012) chama a atenção de seus leitores para um caminho tomado pelo pensamento moderno que precisa ser reconstruído: a fundamentação da filosofia sobre um sujeito completamente solitário. Ela argumenta, contudo, que o filósofo só tem a oportunidade de pensar trancado em seu quarto porque aprendeu socialmente a usar a linguagem, mesma via pela qual aprendeu a pensar⁶⁸.

A autora cita Vygotsky (*apud* CAVELL 2012) acerca do processo de construção do plano interno de um sujeito:

[...] a criança tenta alcançar, sem sucesso, um objeto, e a mãe vem para pegar o objeto para a criança. Com o tempo, aquilo que era gesto primeiro e que não tinha nenhum significado para a criança, mas apenas para a mãe, torna-se, por meio dessa interação, um gesto com o qual a criança significa apontar. Por essa via, eventualmente a criança, assim como a mãe, pode apontar para a maçã com a ideia em mente de “aqui está uma maçã”. (CAVELL, 2012, p. 3)

Ela também menciona o trabalho do segundo Wittgenstein (1953/1979), para quem a linguagem só pode ser construída por meio do engajamento ativo e direto dentro de uma comunidade de falantes. Há maneiras apropriadas ou não de usar um conceito, e isso só pode ser regulado socialmente, por um conjunto de regras compartilhadas onde o indivíduo aprende praticando. Assim, as noções de certo e errado, verdadeiro e falso, bem e mal, eu e o outro e também a noção de mente só podem fazer sentido para o indivíduo a partir de sua participação em uma comunidade.

7. Na *práxis* do uso da linguagem, um parceiro enuncia as palavras, o outro age de acordo com elas; na lição de linguagem, porém, encontrar-se-á este processo: o que aprende *denomina* os objetos. Isto é, fala a palavra, quando o professor aponta para a pedra – Sim, encontrar-se-á aqui o exercício ainda mais simples: o aluno repete a palavra que o professor pronuncia – ambos processos de linguagem semelhantes. (WITGENSTEIN, 1953/1979, p.12)

A noção de triangulação, utilizada por Davidson para explicar a relação de interdependência entre linguagem, mundo e os sujeitos linguísticos vincula a mente humana à realidade social. Nesse sentido, Davidson se considera wittgensteiniano, pois para o Wittgenstein das *Investigações Filosóficas* (1953/1979), a linguagem é social. Assim, Davidson (1990/2001e) argumenta que não é possível sequer ter crenças e nenhum tipo de estado mental proposicional se não há a relação dinâmica que envolve três elementos: o

⁶⁸ Certamente, algumas mudanças no pensamento filosófico foram necessárias para que Marcia Cavell pudesse nos trazer esse argumento. O pensamento moderno, por sua vez, foi importante para que tais mudanças pudessem sobreviver, em especial, o pensamento cartesiano, fundamental para o surgimento da disciplina “Filosofia da Mente”.

sujeito, seu interlocutor e o mundo. Quando aprendemos a falar, os nossos diversos professores estabelecem essas relações com o mundo. Portanto, o fato de a linguagem estar aí para pensarmos sobre a mente traz consigo fortes indícios de que o mundo existe e de que os interlocutores possuem estados mentais apropriados para falar dele.

De acordo com Leclerc (2005, p. 159-160), há três apresentações da triangulação na obra de Davidson. A primeira apresenta duas criaturas carentes de linguagem e de pensamento conceitual reagindo ao mesmo evento e tentando cada uma associar as reações da outra a esse evento. A segunda envolve o aprendiz de uma linguagem em um ambiente social que contém professores. Na terceira, dois falantes de línguas diferentes e que não possuem uma língua em comum interagem. As duas últimas são bastante claras para compreendermos a noção de triangulação, pois a maneira como o aprendiz ou os falantes de línguas diferentes conseguem ajustar suas reações linguísticas à realidade recorre às definições ostensivas, o que evidencia a necessidade de conexões entre o que é dito e o mundo externo, sem as quais não é possível a linguagem⁶⁹.

Alguém precisa, por exemplo, se comunicar com uma pessoa estrangeira que chegou ao Brasil. Nesse caso, as duas pessoas começam a estabelecer conexões entre os sons emitidos por cada uma e o mundo ao redor. O mundo circundante é um parâmetro fundamental sem o qual não será possível a essas pessoas interpretar as crenças uma da outra. Além disso, elas precisam ter como pressuposto que cada uma está fazendo essas conexões entre suas palavras e o mundo, pois a comunicação não será possível se uma das pessoas acreditar que sua interlocutora é louca e tem um discurso totalmente desconectado do mundo.

A individualização das crenças e pensamentos, mas também de significados, conceitos e outros estados mentais, só pode ser concebida a partir de conexões causais sistemáticas na triangulação entre o indivíduo, o outro falante com quem ele interage e objetos ou eventos no mundo (SILVA FILHO, 2007, p. 122).

De acordo com Davidson (1997/2001d), há dois aspectos importantes do pensamento que se explicam a partir da triangulação: primeiro, sua objetividade, pois o pensamento proposicional possui um conteúdo que é verdadeiro ou falso independentemente daquilo que o

⁶⁹ Seria interessante mostrar que os estados mentais dos animais não humanos também possuem características relacionais. O argumento da triangulação é muito promissor nesse sentido e Davidson (1997/2001d) o aplica a animais não humanos, pois a triangulação é anterior à linguagem. Assim, o aspecto relacional dos estados mentais não depende de linguagem convencional. Portanto, pode se aplicar a qualquer tipo de estado mental, independentemente de ser um estado de uma mente humana. Contudo, conforme apontado por Miguens (2006), fica difícil compreender, no pensamento de Davidson, se é possível haver uma mente pré-linguística, ou o que seria tal mente, justamente por causa do holismo do mental. Isso talvez seja uma lacuna importante no pensamento do autor e convém investigar a profundidade dessa lacuna.

sujeito crê; segundo, o conteúdo empírico dos pensamentos acerca do mundo, pois uma vez que o pensamento está vinculado à triangulação, ele tem uma base estabelecida no mundo.

Miguens classifica a noção de triangulação de Davidson em duas categorias: “i) triangulação pré-cognitiva e pré-linguística, (ii) conceitual e linguística. A primeira envolve animais não humanos e crianças, a segunda apenas humanos com domínio de uma linguagem” (2006, p. 102-103). Ela explica que, no pensamento davidsoniano, o primeiro tipo de triangulação é requisito para que o segundo tipo possa surgir. O termo “pensamento”, para Davidson, é reservado aos seres linguísticos, dirige o sujeito ao mundo físico e social por meio da triangulação.

A autora (2006) propõe o estudo dos desenvolvimentos recentes em ciência cognitiva acerca de *atenção conjunta*⁷⁰ como uma forma de suprir o que considera uma lacuna no pensamento de Davidson acerca da triangulação: a passagem da triangulação pré-linguística para a conceitual. Para ela, um problema no pensamento de Davidson é a abordagem do mental como algo exclusivamente linguístico. O primeiro tipo de triangulação, pré-linguística, é muito empobrecido no trabalho de Davidson, o que gera essa lacuna. Segundo ela, o estudo dos processos intermediários de triangulação, anteriores às mentes linguísticas, traz uma compreensão mais profunda e menos exteriorizada do mental que envolve não apenas as atitudes proposicionais, mas também a percepção, um aspecto fundamental que se perde na passagem da primeira para a segunda triangulação⁷¹.

Cavell (2012) explica que, embora os argumentos de Davidson se sustentem sobre bases meramente lógicas e conceituais, há muitos estudos empíricos desenvolvidos por psicólogos que trazem também esse tipo de evidência para apoiar aquilo que o autor diz sobre a linguagem e o pensamento. Ela cita pesquisas empíricas segundo as quais a compreensão dos próprios estados mentais pressupõe a capacidade de interpretar os outros seres linguísticos, e vice-versa. Se ela está correta, há um movimento de vaivém sem o qual não podemos compreender os outros sujeitos linguísticos assim como não podemos compreender a nós mesmos.

⁷⁰ Atenção conjunta consiste na coordenação de mentes (pré-linguísticas) em comunicação acerca do mundo. As crianças pequenas, por volta dos 12 meses começam a apontar para coisas e situações e, antes dessa idade, a dirigir o olhar para o mesmo foco ao qual o olhar de outrem está dirigido. De acordo com Miguens, a atenção conjunta talvez possa representar o passo intermediário entre a triangulação linguística e a pré-linguística. A autora não considera o trabalho desenvolvido até então por Tomasello como capaz de suprir a lacuna, pois Tomasello também apresenta uma descontinuidade entre o pensamento dos animais e o pensamento humano.

⁷¹ Essa é uma questão bastante relevante: se a abordagem do mental é feita exclusivamente a partir das atitudes proposicionais, então como seria possível haver um tipo de triangulação anterior à linguagem? A passagem da cognição não linguística para a cognição linguística não seria abrupta demais? Como compreendê-la?

Davidson argumenta em favor da triangulação em bases estritamente lógicas e conceituais. Mas há uma boa quantidade de evidências para isso em estudos psicológicos sobre os relacionamentos do bebê e da criança. [...] apontar para objetos, brincar juntos, jogar jogos, revezar-se. A condição, novamente, não é apenas que o bebê e seu cuidador possam apontar para o mesmo objeto, mas que eles possam observar o outro fazendo esta referência (CAVELL, 2012, p. 6).

O trabalho de Tomasello (2014) tem características não apenas fundamentadas em pesquisa empírica, mas também em trabalhos desenvolvidos em filosofia da linguagem, inclusive por Davidson. Tomasello é um psicólogo do desenvolvimento que tem se focado nas questões da relação entre cognição e cultura. Para isso, ele costuma usar estudos comparativos entre as crianças pequenas e os grandes símios. Em muitos aspectos, seu trabalho parece bem coordenado com o de Davidson acerca do papel da linguagem na constituição do modo de pensar humano. Porém Tomasello busca respostas também em nossos ancestrais evolutivos, pois o surgimento da linguagem é gradual e depende da existência de estruturas anteriores. Nesse sentido, uma história evolutiva da linguagem e do pensamento pode ser contada de modo a corroborar com a tese da triangulação.

Um último aspecto que vale observar acerca da noção de triangulação é que ela é capaz de trazer uma consistência à mente, pois é uma explicação de como o sujeito internaliza o mundo objetivo. Ocorre que, com a socialização e a aprendizagem da linguagem, aqueles conteúdos são internalizados, passando a constituir o universo mental. Esse universo é dirigido ao mundo, mas é constituído pelas atitudes proposicionais daquele indivíduo. Há, portanto, algo além do comportamento na mente humana, mas não é nenhuma substância ou propriedade que isole o sujeito da realidade externa, é um conjunto de estados mentais relacionados entre si, com o mundo e com os demais sujeitos de uma realidade social.

2.3.1. Pensando em bases evolutivas para a triangulação

Embora os estudos acerca do mental sejam considerados por muita gente algo que sempre se esquivará de qualquer tipo de abordagem científica⁷², o tipo de concepção do mental aqui estudado não está fechado a essa possibilidade. Se a mente é produto das relações do sujeito com o mundo e com os outros sujeitos, se é construída a partir de uma base comum de crenças compartilhadas e se essas crenças são majoritariamente verdadeiras, então as

⁷² Dennett é um defensor da possibilidade de explicação para o mental com bases objetivas. Sobre os que encaram a mente como um grande mistério, ele diz o seguinte: “Nós nunca, jamais, entenderemos a consciência, *eles* declaram; ela vai sistematicamente desafiar os melhores esforços da nossa ciência e filosofia até o fim dos tempos” (DENNETT, 2013, p. 179, grifo nosso, tradução nossa).

relações existentes entre objetivo e subjetivo podem ser esclarecidas sem a necessidade de manter a mente como uma joia misteriosa e inalcançável dentro de um cofre sem chave.

A proposta, aqui, não é de uma redução gananciosa de toda explicação acerca do mental aos métodos e vocabulários da física (como disciplina científica), pois as diversas disciplinas possuem métodos diversificados que podem ser muito mais úteis. Os trabalhos desenvolvidos a partir de estudos científicos sérios, especialmente nas ciências humanas e biológicas, podem contribuir bem mais com as hipóteses filosóficas acerca do mental, corroborando-as ou se apresentando como argumentos divergentes, mas que devem ser enfrentados. Davidson é um autor que segue um método argumentativo filosófico. Seus raciocínios, contudo, também podem ajudar a constituir uma fundamentação teórica para cientistas preocupados em compreender a mente humana⁷³. Há, aqui, uma via de mão dupla.

Muitos trabalhos desenvolvidos na área de psicologia têm contribuído com uma corroboração das ideias de Davidson acerca de triangulação e também da caridade. Nesse sentido, como mencionado, encontramos no livro sobre a evolução da linguagem do psicólogo do desenvolvimento Michael Tomasello (2014) algumas referências a Davidson e é possível observar que a noção de triangulação desempenha um importante papel na fundamentação da pesquisa desenvolvida pelo psicólogo. Seu trabalho se dedica a especificar as características da cognição humana a partir de suas origens evolutivas. Acerca da triangulação, Tomasello a apresenta como algo que direciona os animais nela envolvidos para a objetividade:

[...] o receptor, para fazer o salto abduutivo necessário para alcançar as intenções comunicativas do comunicador, tem que então simular a perspectiva dele em sua perspectiva (ao menos). Essas transações em perspectivas significaram que os primeiros indivíduos humanos não apenas representaram o mundo diretamente para si mesmos, assim como todos os macacos mas, adicionalmente, ao menos em alguns aspectos, experimentaram exatamente o mesmo mundo visto simultaneamente de diferentes perspectivas sociais. Esse processo de triangulação inseriu pela primeira vez um pequeno, mas poderoso rasgo entre o que nós poderíamos agora chamar de objetivo e subjetivo (TOMASELLO, 2014, p. 70, tradução nossa).

Para que a forma de pensar e agir propriamente humana pudesse surgir, segundo Tomasello, foi necessário haver antes relações de triangulação entre primatas não humanos e mundo. Essa hipótese já havia sido defendida por Davidson sobre bases filosóficas. Tentemos

⁷³ Davidson é, na realidade, um autor muito influente. “A extensão e unidade de seu pensamento, em combinação com o caráter às vezes conciso de sua prosa, significa que Davidson não é um escritor fácil de abordar. Todavia, por mais que seu trabalho possa por vezes parecer exigente, isso de forma alguma diminui a importância deste ou a influência que tem exercido e indubitavelmente continuará exercendo. Na verdade, nas mãos de Richard Rorty e outros, e através da difundida tradução de seus escritos, as ideias de Davidson alcançaram uma audiência que se estende para muito além dos confins da filosofia analítica de língua inglesa. Dos filósofos americanos do final do século XX, talvez apenas Quine tenha tido receptividade e influência similares” (MALPAS, 2015, tradução nossa).

compreender um pouco essa história, pois talvez ela traga fundamentos empíricos e evolutivos tanto para a tese davidsoniana da triangulação quanto para o princípio de caridade. Essa relação entre o surgimento do pensamento objetivo e a triangulação é importante, ancorando algumas de nossas crenças no mundo e trazendo um lastro para o conjunto total das crenças das pessoas. Nesse sentido, afirma Davidson:

A identificação dos objetos do pensamento repousa sobre uma base social. Sem uma criatura que observe outra, não teria lugar a triangulação que situa os objetos relevantes em um em um espaço público. Com isso não quero dizer que uma criatura observando a outra proporcione o conceito de objetividade a qualquer uma delas; a presença de uma ou mais criaturas interagindo uma com a outra e com o entorno comum é, na melhor das hipóteses, uma condição necessária para tal conceito. Apenas a comunicação pode proporcionar o conceito, pois ter o conceito de objetividade, os conceitos dos objetos e eventos que ocupam um mundo compartilhado, de objetos e eventos cujas propriedades e existência é independente do nosso pensamento, requer que estejamos informados do fato de que compartilhamos pensamentos e um mundo com os demais (DAVIDSON, 1990/2001e, p.202, tradução nossa).

Na concepção de Tomasello, há algo que faz a cognição humana única, é o pensamento cooperativo (antes da linguagem). Ele defende a hipótese da intencionalidade compartilhada, segundo a qual as pressões seletivas forçaram as atividades compartilhadas, criando condições para o desenvolvimento da linguagem e da cultura institucionalizadas. Além disso, a compreensão do surgimento da linguagem, na concepção de Tomasello, passa pela capacidade de estabelecer relações cooperativas que envolvem o sujeito, ao menos um parceiro comunicativo e o mundo.

Como vimos, Davidson (2001d, por exemplo) também considera a cognição humana muito específica. A causa, para ele, é a existência da linguagem complexa em nossa espécie. Ao tentar especular sobre o surgimento do pensamento, ele considera essa tarefa muito difícil, pois o pensamento não pode ser tratado isoladamente de uma série de outros conceitos igualmente complexos, como crença, desejo, intenção, ação, consistência, racionalidade (aqui, acrescentaríamos também a linguagem como um conceito que, na concepção de Davidson, está imbricado ao de pensamento⁷⁴)... Estamos diante do holismo do mental: a ideia de que os termos mentais não podem ser pensados fora da sua relação com outros conceitos mentais.

Em *The Second Person* (1992/2001c), Davidson apresenta o caráter abstrato da linguagem. “[A]s próprias expressões continuariam sendo abstratas e sua existência, independente da exemplificação” (DAVISON, 1992/2001c, p. 107, tradução nossa). Porém, há as atividades linguísticas reais: os discursos falados e escritos pelas pessoas em nossa vida

⁷⁴ Na visão de Tomasello, o pensamento é anterior à linguagem, mas a linguagem molda um modo diferente de pensar, especificamente humano.

quotidiana. Essas atividades linguísticas são compreendidas por meio da interpretação. Aqui, Davidson defende que a noção de falar uma língua requer que “cada falante intencionalmente se faça interpretável para o outro” (1992/2001c, p. 115). Isso envolve, tanto a auto atribuição de atitudes proposicionais como sua atribuição ao interlocutor. Conseguir dar significado a um discurso, nessa concepção, é obter êxito na tarefa de ser interpretado, o que inclui atribuição mútua por parte de ao menos duas pessoas de atitudes proposicionais umas às outras, vinculando esses estados mentais ao mundo circundante. Assim, forma-se um triângulo entre falante, intérprete e mundo que é, na concepção de Davidson, fundamental para que exista linguagem. O mesmo se estende às atitudes proposicionais. Segundo o autor, elas também requerem a existência da triangulação.

Porém, antes do desenvolvimento da razão, houve estágios nos quais os conceitos de crença, desejo, intencionalidade e os demais conceitos mentais não existiam. Isso ocorre tanto na história evolutiva humana (filogenia) quanto na história do desenvolvimento de cada ser humano (ontogenia). Embora as referências de Davidson à evolução sejam escassas, elas demonstram que o autor está ciente da abertura existente entre seu trabalho e o pensamento evolutivo.

O pensamento, para Davidson, é proposicional. Nesse sentido, a acusação de antropocentrismo⁷⁵ acaba por lhe caber. Tomasello, diferentemente, atribui pensamento já aos grandes símios. Há três processos cognitivos que, na concepção deste, são considerados suficientes para que uma criatura seja considerada pensante: (1) representação cognitiva de experiências não presentes temporal nem espacialmente (quer dizer, experiências que não estão ocorrendo naquele momento diante do sujeito podem ser representadas por ele); (2) habilidade de fazer simulações e inferências protológicas, transformando essas representações para prever seus desdobramentos; (3) automonitoramento, isto é, avaliação dos resultados das decisões comportamentais, com tomada de novas decisões a partir desses resultados. Os grandes símios, segundo ele, possuem todas essas habilidades. Não lhes falta o pensamento. O que lhes falta é o pensamento objetivo, reflexivo e normativo que caracteriza a razão. Podemos considerar a própria construção da objetividade do pensamento como algo enraizado em nossa história evolutiva. No caso do ser humano moderno, ela depende da vida em sociedade, mas a vida em sociedade está apoiada pelos nossos ancestrais.

⁷⁵ Mencionada em seção anterior. Convém notar, contudo, que a acusação de antropocentrismo feita por TENNANT (1999) não se direciona à defesa da existência de pensamento nos animais não humanos, mas sim à possibilidade lógica de que algum outro ser, inclusive extra terrestre, possa pensar. De todo modo, dadas as habilidades cognitivas comportamentais apresentadas por alguns animais não humanos, não parece haver razões para restringir ao grupo dos humanos a possibilidade de pensamento. Talvez isso não represente um abalo nos pilares do pensamento davidsoniano e a situação possa ser resolvida por meio de um ajuste terminológico.

O ser humano moderno, segundo ele, possui um tipo de intencionalidade própria: a intencionalidade coletiva. Esse tipo de intencionalidade se volta para um modo de pensar cultural característico de uma comunidade. Os humanos modernos precisam se coordenar com qualquer outra pessoa, de dentro e de fora de seu grupo social. O pensamento passa a se compor de representações de uma realidade social repleta de elementos abstratos culturalmente constituídos (como os papéis individuais que cada um exerce no grupo, o dinheiro, as leis...). A linguagem também ganha uma dimensão objetiva, por se conectar aos outros seres humanos e ao mundo. Esse modo de pensar está vinculado à linguagem articulada e socialmente compartilhada. Para que isso pudesse acontecer, na concepção do autor, a intencionalidade passa por dois estágios anteriores: o estágio da intencionalidade individual, presente também nos grandes símios e o estágio da intencionalidade compartilhada, após a separação dos ramos da árvore⁷⁶.

Qualquer tipo de intencionalidade apresentado por Tomasello se desenvolve no contexto da triangulação⁷⁷. Porém, os dois tipos mais recentes envolvem colaboração. A intencionalidade compartilhada se torna presente em alguma espécie homínida que nos precedeu. Tomasello supõe que tenha surgido em um ancestral comum entre nós e os Neandertais, há 400 mil anos: o *Homo heidelbergensis*, mas determinar o momento histórico do surgimento da colaboração e da intencionalidade compartilhada não é ponto central. O importante é pensar nos pressupostos do pensamento proposicional, pois assim compreendemos a base comum que conecta nossos conteúdos mentais ao mundo compartilhado e ao outros sujeitos.

Os grandes símios, conforme os resultados dos experimentos estudados pelo autor, possuem boa habilidade para o uso de ferramentas, o que sugere a existência de simulações mentais e imaginação desenvolvidas o suficiente para se projetarem temporal e espacialmente. Além disso, eles têm habilidades procedimentais que envolvem inferências protológicas, o que é demonstrado por um experimento de Call (2004, *apud* Tomasello 2014) no qual dois copos são apresentados para chimpanzés, sendo que um copo contém um alimento dentro e o outro está vazio. O experimentador, então, sacode um dos copos perante o chimpanzé. Se o copo sacudido fizer barulho, o chimpanzé o escolhe. Caso contrário, escolhe o outro copo.

⁷⁶ Compartilhamos muito do nosso DNA com os grandes símios. Pode-se presumir que as características cognitivas presentes em ambos já existiam em um ancestral comum. O ancestral comum mais recente entre nós e eles marca a divisão da árvore evolutiva em dois ramos, sendo que nós pertencemos a um e os grandes símios pertencem ao outro. Assim, podemos presumir também que as habilidades cognitivas não compartilhadas foram desenvolvidas nos ramos específicos.

⁷⁷ Tomasello não apresenta uma concepção de intencionalidade tão ampla quanto a de Dennett. Como veremos, para este a intencionalidade está presente até mesmo na relação entre chave e fechadura.

Segundo Tomasello, algumas noções protológicas são necessárias para que o chimpanzé consiga realizar essas tarefas. Ele precisa ter uma noção pré-proposicional de negação (o copo que se mantém silencioso quando sacudido *não* contém o alimento) e de consequência (se o copo faz barulho, ele contém um alimento), além de precisar fazer inferências a partir daí e tomar decisões comportamentais fundadas nessas inferências⁷⁸.

As habilidades desenvolvidas pelos grandes símios foram no contexto da competição por alimentos, parceiros e outros recursos importantes dentro do grupo social. Com isso, passam a reconhecer indivíduos, formar relações hierárquicas e reconhecer papéis dentro do grupo, além de prever os comportamentos alheios. “Os grandes símios não humanos não apenas são agentes intencionais como compreendem os outros como agentes intencionais” (TOMASELLO, 2014, p.20, tradução nossa).

Os experimentos mostraram capacidade nos grandes símios de manipulação do comportamento alheio por gestos de apontar coisas no mundo para chamar a atenção de outro. Isso sugere uma forma de comunicação desenvolvida a partir de uma relação em que todos os envolvidos no processo compartilham de uma base comum ancorada no mundo. Ao apontar para algo, o símio sabe que o outro perceberá aquilo e que ambos compartilham de alguma percepção comum. O comportamento deles, segundo Tomasello, indica uma orientação à competição porque se dirige à busca de recompensas individuais. Não demonstram comportamentos voltados para o benefício exclusivo do outro, como nós demonstramos. Convém destacar, aqui, que se o processo cognitivo se ancora em uma base de conhecimentos compartilhados, isso fundamenta o princípio de caridade em Davidson.

Outro aspecto importante apontado por Tomasello em relação à cognição dos grandes símios é que eles monitoram suas ações com vistas a tomar boas decisões em ocasiões futuras sendo capazes, por exemplo, de fazer algo desagradável visando uma recompensa agradável. Além disso, demonstram avaliação de seus estados psicológicos, pois se comportam de modo diferente, por exemplo, quando sabem onde o experimentador escondeu a comida e quando não sabem. Todas essas observações acerca do comportamento dos grandes símios levam Tomasello a defender que eles pensam, pois possuem as características requeridas: representações cognitivas abstratas; capacidades protológicas de realizar inferências e

⁷⁸ Ora, se os grandes símios já apresentam essa forma protológica de se comportar, é o que nos basta para aceitarmos que nosso interlocutor humano, seja ele quem for, compartilha conosco de um amplo fundamento lógico comum, ainda que pertença a uma cultura que não tenha desenvolvido uma simbolização escrita para a lógica, obviamente. Assim, a tese de Quine de que não há entre seres racionais uma mentalidade pré-lógica é corroborada, pois um olhar protológico para a realidade que nos cerca já estaria presente em um ancestral antes do desenvolvimento da linguagem tipicamente humana.

automonitoramento psicológico. O pensamento, segundo essa concepção, é independente da existência da linguagem e da cultura (ao contrário do que defende Davidson).

Antes de haver a forma de intencionalidade coletiva apresentada por nós, humanos modernos, Tomasello considera necessário haver uma etapa anterior: a da intencionalidade compartilhada. É preciso haver em homínídeos ancestrais alguma forma de relação social baseada na cooperação, mas anterior ao desenvolvimento da linguagem convencional e da cultura cumulativa. Provavelmente, o aumento das populações e a escassez de alimentos tenha provocado a necessidade de forrageamento colaborativo, sem o qual a aquisição de alimentos teria se tornado impossível. Pela mesma razão, deve ter surgido competição entre os grupos de homínídeos, gerando a necessidade de colaboração entre os membros de dentro de um grupo. De acordo com o psicólogo, as pesquisas mostram que onde se observa competição entre os macacos, observa-se colaboração entre os humanos.

Por exemplo, segundo Tomasello (2014, p.34), os chimpanzés também se engajam em atividades cooperativas na aquisição de alimentos. Porém, quando lhes é dada uma oportunidade de escolher, preferem tentar adquirir alimentos sozinhos, provavelmente porque assim também poderão usufruir sozinhos de sua recompensa. Além disso, contrariamente a nós, eles preferem se alimentar sozinhos a compartilhar as refeições. Por outro lado, as comunidades humanas tendem a agir de forma colaborativa em tudo: na divisão das refeições, na criação e na educação das crianças, na comunicação e nas estruturas sociais diversas.

Se as pressões ambientais forçaram nossos ancestrais à ação colaborativa para que pudessem conseguir alimentos e competir com outros grupos, então escolher parceiros colaborativos se torna vital, assim como ser escolhido por bons parceiros. Isso, segundo Tomasello, contribui para o desenvolvimento da capacidade de leitura de mentes recursiva. Para o sujeito colaborar, ele crê que o outro espere isso dele. A base comum de conhecimentos compartilhados se amplia e, a partir de então, passa a envolver também o conhecimento que cada parte tem daquilo que a outra parte busca alcançar. Outra habilidade cognitiva importante também surge: o automonitoramento social. Esses homínídeos tiveram que desenvolver uma preocupação com a maneira como os outros indivíduos os avaliam enquanto parceiros. A avaliação social que os outros fazem do meu comportamento, assim, se torna algo muito relevante.

Esses homínídeos que precisaram lidar com as primeiras formas de atividade colaborativa agiam sobre aspectos compartilhados do mundo. É assim que a intencionalidade passa a ser compartilhada. Só há intencionalidade compartilhada quando ao menos dois sujeitos se voltam para algo disponível a ambos, mas isso não é suficiente. Cada um deles

precisa *saber* que o outro também está voltado para aquele objeto. É preciso, aqui, também atribuir ao outro a concordância em relação ao agir compartilhado. Nesse sentido, os papéis e as perspectivas individuais se tornam cruciais para que o objetivo comum possa ser alcançado. Note-se que, segundo Tomasello, tudo isso é feito pelos homínidos de modo não linguístico, pois a linguagem articulada ainda não havia se desenvolvido.

2.3.2. *Intencionalidade compartilhada e linguagem*

O que aparece nesses homínidos que nos antecederam é chamado por Tomasello de engajamento de segunda pessoa. Nesse momento, ao menos dois indivíduos passaram a se envolver com os estados intencionais dos outros de maneira compartilhada e recursiva. Isso fundamenta o passo seguinte da evolução, onde a linguagem convencional e a cultura cumulativa se tornam possíveis.

Junto da colaboração, segundo Tomasello, surgiu a comunicação por gestos e mímicas, levando a uma nova forma de pensamento. Essa comunicação se torna importante porque, para cada parceiro, é interessante contribuir para que o outro consiga desempenhar seu papel na busca de algo que interesse a ambos. O autor considera apontar e fazer mímicas como gestos universais humanos que independem do desenvolvimento de uma linguagem convencional. Estão fundados na triangulação e na base cognitiva comum entre as pessoas envolvidas. “Se eu aponto para você em direção a uma árvore ou represento gestualmente [“pantomime”] para você uma árvore, sem nenhum solo comum, você não tem nada que possa guiar suas inferências sobre o que eu pretendo lhe comunicar e por que” (TOMASELLO, 2014, p. 50, tradução nossa). Assim, surge também o pressuposto de que o outro quer comunicar algo relevante para o sujeito.

[...] o que nós temos neste ponto na nossa história evolutiva da comunicação humana são tentativas individuais para coordenar seus estados intencionais, assim como suas ações, apontando situações novas e relevantes um para o outro. Isso depende de eles terem certo montante e tipo de solo comum e requer, no mais, que os interagentes façam uma série de inferências sociais recursivas sobre as perspectivas e estados intencionais de um e do outro (TOMASELLO, 2014, p. 59, tradução nossa).

As mímicas são um modo de comunicação ainda mais poderoso, pois direcionam a imaginação do outro a entidades ou a situações não presentes. De acordo com Tomasello, nenhum primata não humano é capaz de usar gestos para simbolizar algo ausente justamente porque não entendem a comunicação fundamentada na cooperação. Além disso, uma vez que

há várias formas diferentes para representar gestualmente coisas e eventos diversos, a mímica é considerada pelo autor como o rudimento da semântica, e tem potencial categórico. A motivação da comunicação pela mímica é colaborativa. Pressupõe-se que o outro quer dizer algo que interessa ao sujeito.

O psicólogo argumenta que a combinação entre gestos é consequência natural na evolução da comunicação. Participantes e situações, assim como sujeitos e predicados, passam a receber representações gestuais. Aos poucos, torna-se possível, inclusive, representar por gestos aquilo que é impossível no mundo material. “[...] eu poderia produzir um gesto icônico para viajar seguido de um apontamento para um local, o qual eu generalizo para qualquer local. Eu poderia então imaginar ou comunicar, via esse esquema, nossa criança viajando para o sol” (TOMASELLO, 2014, p. 73, tradução nossa). A comunicação, segundo ele, se torna *quase proposicional* e surge a necessidade de se considerar a perspectiva do outro sobre a mesma realidade (aqui estão os primórdios da perspectiva objetiva). O comprometimento com a verdade se torna mais e mais importante. Essa etapa teria sido necessária para que pudesse surgir a comunicação baseada em convenções⁷⁹.

A linguagem convencional estrutura a mente do humano moderno com atitudes proposicionais. Agora, é possível diferenciar minhas crenças e as crenças dos outros, assim como as crenças verdadeiras e as falsas. Portanto, a existência das crenças se funda sobre o surgimento da objetividade das normas e instituições que constituem a realidade culturalmente compartilhada. A disposição para punir alguém que tenha cometido ato injusto contra terceiro, segundo Tomasello, não é uma característica comportamental que possa ser observada nos grandes símios, ao passo que as crianças de três anos já usam uma linguagem normativa objetiva para tratar de situações nas quais não estão envolvidas. Esses padrões e normas são ensinados e aprendidos, formam um pano de fundo objetivo contra o qual os indivíduos são avaliados e julgados. Assim, a nova forma de cognição envolve o automonitoramento normativo, as ações do sujeito são reguladas pelas normas do grupo. Após a linguagem se tornar convencional, o pensamento humano ganha recursos que só podem existir em um ambiente cultural.

⁷⁹ Várias espécies de animais possuem aprendizagem social. Entre os chimpanzés e bonobos, observam-se comportamentos que são específicos de determinados grupos, que persistem temporalmente e requerem aprendizagem social. Esses animais observam comportamentos em outros e aprendem. Conforme Blackmore (1999), a aprendizagem social ocorre quando há influência da observação ou interação com outro animal na aprendizagem. Em geral, é reforçada por um estímulo ambiental. Há algumas diferenças em relação à aprendizagem humana. Primeiro as crianças humanas costumam imitar detalhadamente o modo como um adulto executa uma tarefa. O foco está no processo, e não no resultado (CALL; TOMASELLO, 1995). Segundo, que os outros animais aprendem por observação e têm o comportamento reforçado por estímulo ambiental, ao passo que os adultos humanos se envolvem ativamente na tarefa colaborativa de ensinar (TOMASELLO, 2014).

O humano moderno nasce em um ambiente no qual as convenções comunicativas estão estruturadas por meio de regras. A escolha de vocalizações é arbitrária e se torna convenção entre um grupo, dando origem às abstrações. Torna-se possível, além disso, conceitualizar situações complexas que envolvem diversos agentes, por exemplo, uma audiência judicial ou um funeral. Note-se, ainda, que a compreensão desse novo tipo de linguagem envolve uma série de inferências sem as quais não seria possível compreender a linguagem. Se eu digo, por exemplo, que vi você andando com uma mulher hoje pela manhã, nós dois saberemos que essa mulher é alguém que eu não conheço e, provavelmente, haverá também o pressuposto que você foi visto andando somente com essa mulher.

Da mesma forma, qualquer compreensão linguística exige uma série de conhecimentos relacionados sem os quais não será possível compreender o que está sendo dito. No mais, falar e esclarecer aquilo que se pretende dizer permite também ao sujeito autorreflexão e automonitoramento a partir da base de conhecimentos comum, às vezes provocando uma necessidade também de esclarecimento dessa base de conhecimentos. Isso leva o ser humano moderno a uma autorreflexão totalmente baseada em normatividade.

2.3.3. Razões e verdade

Os primeiros hominídeos, segundo Tomasello, provavelmente apontaram para pegadas na hora de tomar uma decisão compartilhada. Supondo que dois deles estivessem envolvidos com a caça de um animal e cada qual quisesse ir numa direção, apontar para as pegadas serve como uma *razão* para decidir em qual direção se deve ir, com base nos conhecimentos sobre o mundo que já são compartilhados entre eles. O ser humano moderno lida com razões no sentido completo do termo. As razões permitem explicar para os outros e para si mesmo o que se está pensando e o porquê. O pensamento do humano moderno, por ser linguístico e por ter se construído a partir do contato direto com o mundo, não é privado. Pertence ao mundo do que é compartilhado. A verdade se torna algo relevante para os hominídeos que compartilham, pois se torna um árbitro neutro para fundamentar as decisões compartilhadas. Encontrar junto ao seu companheiro boas razões para agir de uma determinada maneira pode ajudar ambos a encontrar comida e a se livrar de uma situação arriscada. É melhor ter comida do que estar com a razão, por isso, os bons argumentos se tornam importantes, assim como se torna importante convencer o outro quando se crê estar com a razão. Isso encaminha as razões para uma maior objetividade, e o desenvolvimento da linguagem simbólica carrega consigo esse germe.

Na concepção de Tomasello (2014, p. 116), enquanto as inferências dos grandes símios são apenas causais e intencionais, as inferências feitas pelos primeiros humanos são estruturadas recursivamente para interpretação cooperativa. Aos poucos, as razões e justificativas passam a conectar as várias conceitualizações do indivíduo em uma rede inferencial. Os pensamentos proposicionais passam a dar justificção uns aos outros. Criam-se, assim, relações hierárquicas entre as classes das convenções linguísticas. Por exemplo, para alguém saber o que é uma jaguatirica, precisa saber que é um animal, isso permite a construção de uma lógica formal intuitiva, como a que é pressuposta no *princípio de caridade*. As inferências racionais e reflexivas dos humanos modernos se dão no contexto dos padrões normativos de seu grupo cultural. Assim, como defenderia Davidson, o pensamento proposicional se articula de maneira holística.

Os primeiros humanos, na concepção de Tomasello, se envolviam com automonitoramento cooperativo (a partir da perspectiva de um parceiro específico) e automonitoramento comunicativo (antecipação da interpretação de parceiros específicos). Nos humanos modernos, o processo de decisão comportamental é regulado por normas coletivamente aceitas do grupo cultural. Há uma pressão normativa cultural sobre o comportamento do indivíduo. Se eu quero ser membro de um grupo, preciso me adequar às suas normas. É preciso pensar como o grupo. Com isso, a normatividade coletiva é internalizada. O pensamento do sujeito é avaliado pelos padrões normativos do grupo, dando padrões objetivos para o julgamento. Agora, é preciso saber qual é a melhor crença para o grupo, e não para mim. O pensamento se torna coletivo, objetivo, reflexivo e normativo. Essa forma de pensar, segundo Tomasello, caracteriza o raciocínio tipicamente humano.

Ressalta Tomasello que os grupos humanos, ao contrário dos grandes símios, ocuparam todas as partes do globo graças ao surgimento da cultura cumulativa derivada do efeito catraca [“ratchet effect”] (TENNIE; CALL; TOMASELLO, 2009). Tal efeito faz com que o processo de acumulação cultural seja irreversível. As práticas culturais de um grupo são repetidas pelas gerações seguintes até que sejam alteradas, quando então é essa alteração que passa a ser seguida. Em geral, as alterações envolvem um incremento ou melhoria em relação ao que havia antes, gerando a acumulação cultural. O que permite o efeito-catraca é a fidelidade com que as práticas culturais são transmitidas entre os humanos. “A evolução cultural cumulativa entra em cena quando as invenções em um grupo cultural são passadas com tal fidelidade que permanecem estáveis no grupo até que uma invenção nova e melhorada surja (o assim chamado efeito-catraca [...])” (TOMASELLO, 2014, p.121, tradução nossa).

Enquanto alguns autores consideram a linguagem e a cultura cumulativas como fundamentos da cognição humana, Tomasello discorda. Para este, a linguagem, a cultura e o pensamento proposicional são, na realidade, resultado do desenvolvimento de uma nova forma de cognição: essa que envolve a colaboração, a comunicação cooperativa e a intencionalidade compartilhada. Antes de haver linguagem convencional, segundo Tomasello, é preciso haver um modo de pensar fundamentalmente cooperativo. Porém, após o desenvolvimento da linguagem e da cultura, estabelecem-se as bases para reestruturação da cognição, permitindo esse modo de pensar do humano moderno – um modo de pensar que só floresce no solo da cultura.

O solo de conhecimentos comuns se amplia, incluindo conceitos abstratos, de modo que os processos psicológicos e as inferências lógicas se tornem explícitos por meio das atitudes proposicionais, possibilitando pensar sobre o próprio pensamento. Os grupos passam a construir uma racionalidade própria necessária para que possam fornecer razões e convencer seus pares de determinados planos de ação. Essa racionalidade de grupo é internalizada e permite ao indivíduo saber por que ele pensa de determinada maneira e criar uma rede holística de estados mentais proposicionais. Cada indivíduo passa a fazer um automonitoramento de suas ações a partir dos padrões normativos do grupo. Assim, a intencionalidade coletiva cria o pensamento objetivo/normativo.

Embora concorde com muitos filósofos no sentido de defender uma descontinuidade entre as habilidades cognitivas humanas e as dos animais não humanos, há uma discordância entre Tomasello e Davidson, pois para este o pensamento só se desenvolve dentro do contexto da linguagem. Tomasello defende que os animais não humanos possuem pensamento, mas não com as características proposicionais do pensamento humano. O autor propõe a necessidade de um passo intermediário entre os humanos cooperativos e os símios competitivos. “Esse passo intermediário não resolve o problema de Davidson (1982)⁸⁰ de um vocabulário teórico comum que leve do ‘não pensamento’ para o ‘pensamento’, mas ele estreita significativamente a distância a ser atravessada por qualquer passo” (TOMASELLO, 2014, p. 151, aspas duplas do original substituídas por aspas simples, tradução nossa).

Aparentemente, contudo, essa divergência entre os dois autores não é conceitual, mas sim terminológica. Para Davidson, ter pensamento é ter atitudes proposicionais o que, na concepção dele, depende de linguagem⁸¹. Tomasello, por outro lado, defende que o

⁸⁰ Trata-se aqui do artigo “The emergence of thought” (2001d), publicado originalmente em 1997.

⁸¹ A dificuldade em se explicar a emergência do pensamento, para Davidson, é conceitual. O argumento é simples: antes de haver linguagem, para Davidson, não há pensamento. Sem pensamentos, não há conceitos que permitam essa explicação.

pensamento possui essas três características já descritas: representação cognitiva, capacidade de fazer inferência e automonitoramento comportamental. Importante notar, porém, que Tomasello considera a forma de pensamento caracteristicamente humana algo que só se constrói no contexto da cultura e da linguagem convencional. Essa forma de pensar cria uma rede holística de estados mentais que dá origem à objetividade e às razões no sentido completo desses termos: um sentido constituído não apenas de um sujeito em interação com o mundo, mas de vários sujeitos em interação uns com os outros e com o mundo.

Para Tomasello, muitos dos problemas filosóficos surgem precisamente quando nós tentamos entender as coisas abstratamente, fora do contexto evolutivo e funcional. Embora boa parte do pensamento humano pareça inútil, ele foi selecionado evolutivamente pelo seu papel em organizar e regular as ações adaptativas. Para compreender isso, é preciso identificar os problemas relevantes. Tomasello defende que todos os aspectos da cognição especificamente humana são constituídos socialmente, mas há outros aspectos da cognição humana que não dependem da inserção do indivíduo em um contexto social.

2.4. O sujeito intersubjetivo

No pensamento de Davidson, observa-se que as realidades objetiva e subjetiva se constroem uma à outra, sustentam-se mutuamente nessa construção, tornando-se quase transparentes de ambos os lados. Não é que a mente construa o mundo, mas ela é parte desse mundo e se constrói a partir das conexões com ele. A própria concepção do erro se enraíza no fato de que as nossas crenças são majoritariamente verdadeiras. Até a explicação do erro é dada a partir da relação entre o falante, o interlocutor e o mundo objetivo.

Uma criatura que tem conceitos, e, portanto, crenças, irá sem dúvida ter também outras atitudes acerca dos conteúdos proposicionais, tais como intenções, percepções, memórias, desejos, esperanças e o resto. Tal criatura pode ser surpreendida, no sentido de subitamente descobrir que algo que ela acreditava é falso. Com os conceitos vêm as crenças, e com as crenças vem a distinção entre o verdadeiro e o falso. Ter essa distinção é ter o conceito de *objetividade*, ou seja, apreciar o fato de que muitas coisas são como são por mais que pensemos sobre elas. Não se pode ter crenças sobre a maior parte dos aspectos do mundo sem compreender o fato de que as coisas podem ser vistas como, afigurar ou parecer serem outras do que são (DAVIDSON, 1999, p. 40, tradução nossa).

Ora, se um discurso pode ser interpretado a partir de evidências observáveis e assim o significado se revela, revelando também os conteúdos proposicionais do sujeito, a subjetividade deixa de ser algo encarcerado dentro da mente. Agora, temos um caminho que parece conduzir a uma compreensão de como o subjetivo, assim como o significado, podem

se descobrir a partir de evidências publicamente disponíveis. “Eu considero uma teoria do significado como uma parte que não pode ser destacada de uma teoria mais geral do comportamento humano” (DAVIDSON, 1994a, p. 127, tradução nossa). A perspectiva do intérprete, então, tem uma importante função no pensamento de Davidson, pois permite compreender a construção da subjetividade.

Embora o intérprete radical não possa ter como dados iniciais informações detalhadas acerca das atitudes proposicionais do falante, ele precisará partir da atribuição bem geral de alguns estados mentais, como “isso é tomado como verdadeiro” ou “isso é tomado como falso pelo falante”. A compreensão do significado de um discurso, portanto, anda de mãos dadas com a descoberta do que é tomado como verdade naquele discurso. A noção de verdade, então, ganha um papel primordial. É preciso haver algum acordo prévio entre falante e intérprete que envolva, por parte do intérprete, a atribuição de uma série de atitudes ao falante, como racionalidade e sinceridade. A partir daí, o que há na mente do falante pode se revelar concomitantemente ao processo de interpretação e ela se revela apenas a partir daquilo que está publicamente disponível a qualquer intérprete: sons proferidos, comportamento não verbal e mundo.

Outro aspecto importante no processo de interpretação radical é que o intérprete considere que a verdade de algumas crenças do falante esteja apoiada sobre a verdade de outras. O tipo de acesso que o falante tem ao entorno também deve ser levado em consideração pelo intérprete. Isso porque basear o processo de interpretação no mundo circundante acaba sendo frágil. A história da filosofia racionalista vem sempre nos lembrar de que os sentidos são enganosos. É por isso que o intérprete precisa de vários pontos de apoio e de sustentar uns pontos sobre os outros, formando uma rede de crenças relacionadas que se tornam firmes por pertencerem a uma teia forte onde a maior parte das fibras se constituem de verdade.

Nesse sentido, Davidson abraça a ideia de que uma teoria da interpretação não se constrói independentemente das noções de significado e crença. “[A]o interpretar do princípio [“from scratch”] – na interpretação radical – nós devemos de alguma forma oferecer simultaneamente uma teoria da crença e uma teoria do significado” (1975/1984d, p.144, tradução nossa). A relação mútua entre crença e significado faz com que não seja possível abordar essas duas noções independentemente.

Importante observar que a linguagem convencional é uma das instituições humanas. Para ser aprendida por um indivíduo, demora alguns anos e depende de um aparato cognitivo específico. Assim, antes de ter linguagem, cada um de nós já tinha estados mentais

pré-linguísticos. Do mesmo modo, os animais não humanos também possuem estados mentais sem que tenham linguagem.

A abordagem de Davidson acerca do mental não se aprofunda nesses estados mentais não linguísticos, mas é possível que uma abordagem de fundamentação evolucionista possa nos trazer uma compreensão de como é possível, no contato direto com o mundo, passar de mentes não linguísticas para mentes linguísticas. O trabalho de Michael Tomasello não parece oferecer nenhuma contradição direta ao pensamento de Davidson. Ao contrário, traz um apoio para o trabalho do filósofo a partir de experimentos feitos com os grandes símios e as crianças em idade pré-linguística. Porém, o seu trabalho traz pistas para que possamos compreender a mente em geral, e não apenas a mente linguística, a partir da triangulação.

3. INTERPRETIVISMO DE DENNETT: INTENCIONALIDADE ATRIBUÍDA

Um dos temas recorrentes na obra de Dennett é a consciência, o estudo daquilo que aparece para o sujeito a partir de seu próprio ponto vista, isto é, a dimensão de primeira pessoa. Porém, Dennett deseja explicar a consciência de uma maneira compatível com as ciências. Ele necessita, portanto, de uma explicação objetiva. Para isso, recorre à noção de interpretação (fundamentada na noção de tradução de Quine). Um sistema é consciente na medida em que determinada postura precisa ser usada para interpretá-lo: a postura intencional. Nesse sentido, Miguens (2002, p. 24) compreende a teoria da mente de Dennett como normativa ou transcendental. Em outras palavras, para que algo seja compreendido como mental, precisa ser interpretado como tal.

A tarefa de interpretar é considerada por Dennett de extrema importância, pois é a partir dela que nós, humanos, nos tornamos capazes de explicar e prever o comportamento dos outros humanos. Para isso, adotamos a *postura intencional*, a atitude de atribuir estados mentais. O autor considera que essa é uma *habilidade inata*. De fato, os desenvolvimentos em primatologia corroboram com essa tese, como se pode perceber, por exemplo, nas pesquisas desenvolvidas por Tomasello (2014). Se os grandes símios contemporâneos apresentam a habilidade de atribuição de intencionalidade, assim como nós, os humanos, então é possível regredir na cadeia evolutiva, considerando também nosso ancestral comum como portador dessa habilidade.

A postura intencional é uma habilidade considerada importante por Dennett não apenas para nossas relações quotidianas com as outras pessoas. Nós a usamos para compreender também outros animais, assim como uma série de artefatos criados por nós, por exemplo, livros, filmes, computadores e jogos em geral. Também podemos usar essa habilidade para prever o comportamento de coisas mais simples, como um despertador ou mesmo um termostato. Ocorre que há casos nos quais, segundo o autor, não é possível abrir mão da atribuição de intencionalidade: por exemplo, para a compreensão de um discurso, das instituições humanas ou para a compreensão da evolução de estruturas biológicas adaptativas.

A primeira seção deste capítulo trata das diversas posturas que podem ser adotadas para explicação e previsão de sistemas diversos. Em especial, será desenvolvida a noção de postura intencional. A segunda seção trata da noção de psicologia de senso comum, uma

noção bastante central para o interpretivismo de Dennett. Em seguida, o método heterofenômico será apresentado, uma proposta de Dennett para ajudar nas pesquisas sobre a consciência a partir de um ponto de vista de terceira pessoa. A última seção procura mapear diferentes noções de consciência ao longo da obra do autor, procurando colocá-las em diálogo. Há a suspeita de que a memética defendida por Dennett não possa ser conciliada com suas ideias interpretivistas. Esse tema também será tratado⁸².

3.1. Considerações metodológicas

Em geral, a tarefa de interpretar é considerada por Dennett algo muito corriqueiro e natural na vivência humana, mas com especificidades que parecem contradizer as intuições comuns. O autor considera uma prática importante buscar desconstruir as percepções das práticas quotidianas, as quais frequentemente são fruto de hábitos de pensamento profundamente arraigados. Nessa linha, ele faz uma recomendação metodológica aos filósofos e aos leitores de filosofia:

[...] procure por “certamente” [“surely”] no documento e cheque cada ocorrência. Não é sempre, nem na maioria das vezes, mas frequentemente a palavra “certamente” é tão boa quanto uma luz intermitente localizando um ponto fraco no argumento, uma etiqueta de aviso de um suporte de lança de navio [“boom crutch”]⁸³. Por que? Porque marca a beira do que o autor realmente tem certeza e espera que os leitores também terão certeza. (Se o autor tivesse certeza que todos os leitores concordariam, não valeria a pena mencionar.) Estando na beira, o autor teve que tomar uma decisão pessoal sobre o esforço ou não de demonstrar o ponto em questão, ou fornecer evidência para ele e - porque a vida é curta - decidiu em favor da afirmação simples, com a antecipação da concordância presumivelmente bem fundamentada (DENNETT, 2013, p. 53, tradução nossa).

Muitas vezes, tomar como dado aquilo que parece certo ou óbvio, na concepção de Dennett, pode atrapalhar o percurso desde o início, impedindo a superação das ilusões e preconceitos. Por outro lado, perspectivas aparentemente contrárias às intuições podem ser fecundas, desde que cuidadosamente avaliadas. Parte do trabalho para uma explicação da consciência, segundo Dennett, envolve uma compreensão de porque ela nos parece dessa

⁸² Para uma apresentação completa, detalhada e crítica da obra de Dennett desde 1969 até 2001, ver Miguens (2002).

⁸³ Esse termo é usado por Dennett (2013, p. 48-52) como uma metáfora para mostrar uma falha na crítica de Gould ao adaptacionismo. Trata-se de uma espécie de forquilha que apoia uma lança presa a uma base fixa ao mastro do navio na outra extremidade. Essa lança pode se mover ao redor do mastro, em um movimento de compasso. Para o suporte ficar em pé, a lança precisa estar apoiada nele. Dennett, como adaptacionista, pensa que vale a pena buscar uma explicação funcional para uma estrutura biológica, localizando-a no tempo e no espaço, pois as adaptações são locais. Nesse sentido, a rejeição ao adaptacionismo só se manteria em pé quando apoiando os casos não-adaptativos, que são minoria.

forma. Explicando-se isso, o mistério se diluiria. Esse mesmo método, segundo Dennett, poderia ser aplicado de forma ubíqua dentro da metodologia filosófica, trazendo benefícios.

Nesse sentido, ele propõe o estudo do mental pela heterofenomenologia, método para estudar os aspectos subjetivos do mental que consiste na interpretação dos relatos de alguém sobre seus estados internos. “Heterofenomenologia é o estudo dos fenômenos de primeira pessoa a partir do ponto de vista da ciência objetiva” (2013, p. 342, tradução nossa). Na aplicação do método heterofenômico, o sujeito que está sendo estudado contribui por meio da narrativa de suas experiências. A linguagem é o mecanismo que ele tem para tornar públicos seus estados internos. Todos os comportamentos daquele sujeito, contudo, devem ser considerados. A interpretação do discurso desse sujeito constitui o seu mundo heterofenômico.

De acordo com um texto de 1990, o mesmo procedimento interpretativo é adotado ao se interpretar um texto (inclusive de ficção), o discurso de uma pessoa, um artefato cuja função desconhecemos ou uma estrutura biológica (este último exercício de interpretação chamado adaptacionismo). Sendo o mesmo procedimento, a interpretação está sempre sujeita à indeterminação. Em outras palavras, diante de interpretações rivais, não há parâmetros capazes de estabelecer a interpretação correta em nenhum caso.

Segundo Dennett (1990, p. 178), ao interpretar um texto, é comum criarmos hipóteses acerca das intenções do autor. Porém, a interpretação que o autor fornece de seu próprio texto não deve ser considerada a única interpretação correta. No fim das contas, depois de lançar o seu texto, o autor se torna um intérprete entre outros, não há uma interpretação que seja privilegiada ou definitiva. O autor pode, inclusive, confundir-se acerca do que está dizendo. Isso torna a interpretação radicalmente inescrutável, no mesmo sentido em que Quine argumenta pela inescrutabilidade da referência que provoca a indeterminação da interpretação. Essa indeterminação vai muito além da interpretação de textos, ela se estende à interpretação de qualquer comportamento⁸⁴.

Nesse sentido, embora tenhamos uma intuição de que haja algo fixo na mente do sujeito submetido a uma interpretação e que isso poderia resolver os casos de interpretações rivais, Dennett acredita que isso seja apenas uma impressão falsa. Certamente, o sujeito tem um papel importante na interpretação, ao esclarecer aquilo que quer dizer. Porém, lançadas as palavras no mundo, o autor deixa de ser autoridade sobre elas. Nada há dentro do sujeito que

⁸⁴ Davidson (1977/1984f) também enfrenta a questão na inescrutabilidade da referência, mas de modo diferente. Ele propõe que se abra mão da noção de referência em uma teoria do significado, mas sem abrir mão da noção de ontologia.

possa direcionar uma interpretação em detrimento de outras. O mesmo vai ocorrer em relação à interpretação de romances. Não é possível definir a melhor interpretação, ela não existe.

O caso da interpretação dos artefatos pode iluminar a compreensão desse “princípio de não autoridade autoral” (1990, p. 182, tradução nossa). Dennett pede para o leitor imaginar um arqueólogo que encontra um artefato em suas escavações, mas não sabe a utilidade. Ao elaborar uma hipótese de função para aquele objeto, o arqueólogo poderá considerá-la corroborada se o artefato for útil no desempenho da função suposta. Porém, o mesmo objeto poderia ser útil também no desempenho de outras funções imperceptíveis para o arqueólogo. Pode, inclusive, ter sido projetado com a intenção de desempenhar outra função. Ora, mesmo que fosse possível entrar em uma máquina do tempo para descobrir as intenções do projetista, isso não resolveria a questão da indeterminação da interpretação da função daquele objeto, pois qualquer artefato pode mudar de função quando utilizado por alguém. Dennett dá como exemplo o ferro a carvão. Ele foi projetado para passar roupas, mas ninguém mais dá a ele esse uso, melhor desempenhado hoje pelo ferro elétrico. Contudo, muitas pessoas continuam comprando, pois consideram o objeto bonito para segurar portas ou como peso para papéis.

Em relação à interpretação de estruturas biológicas em geral, elas podem ter desempenhado uma função em um determinado local e em um determinado momento, e essa função ter se alterado ao longo da história evolutiva, gerando uma exaptação. Uma exaptação é algo que se desenvolve a partir de pressões evolutivas para desempenhar uma função, mas posteriormente passa a desempenhar outra função. Isso também pode trazer um grau alto de indeterminação em relação à função biológica de um organismo.

De todo modo, ainda que não haja uma base fixa para decidir entre interpretações rivais de nenhum tipo (nem para pessoas, nem para romances, nem para outros artefatos, nem para estruturas biológicas), em todos os casos, na concepção de Dennett, é necessário adotar a *postura intencional* frente ao que quer que se queira interpretar. Será necessário atribuir intencionalidade a alguém ou a algo, nisso consiste o processo de interpretação. No caso dos discursos das pessoas, só poderemos interpretá-los por meio da atribuição de intencionalidade ao autor. O mesmo vale para os romances e demais artefatos. Em relação aos organismos biológicos e suas partes, seria necessário atribuir intencionalidade ao processo de seleção natural. Ao adotar essa postura hermenêutica, conforme o autor, poderemos obter resultados bastante satisfatórios.

Interpretar, portanto, envolve a adoção da postura intencional. Por isso, a aplicação do método heterofenômico pressupõe a adoção dessa postura. A postura intencional, na concepção de Dennett, é algo que fazemos naturalmente, não apenas ao interpretar o discurso

de um sujeito acerca de seus próprios estados mentais, como também para explicar e prever o comportamento de uma série de outras coisas, inclusive as que não são intencionais. A diferença, no fim das contas, é de grau. Atribuímos muito mais intencionalidade quando precisamos interpretar um discurso, afinal, até mesmo para a compreensão de uma simples frase, é necessário fazer uma enorme quantidade de atribuições de crenças e desejos, além da necessidade de atribuir alto grau de racionalidade e coerência ao emissor daquele discurso.

Há várias atitudes possíveis para explicar e prever o comportamento de um sistema qualquer. Essas atitudes vão desde a postura física, na qual o sistema é explicado a partir de sua constituição física, passam pela postura de projeto, na qual é necessário atribuir intencionalidade ao projetista (sendo que essa atribuição de intencionalidade não implica em atribuição de teleologia) e vão até a postura intencional. Quando se trata da aplicação do método heterofenômico, estamos no nível da linguagem. Neste caso específico, só há uma postura capaz de trazer algum sentido ao discurso daquele sujeito acerca de seus estados internos: a postura intencional. Este capítulo trata de duas noções centrais ao interpretivismo de Dennett: a teoria dos sistemas intencionais e o método heterofenômico.

3.1.1. Posturas para explicar e prever o comportamento de um organismo⁸⁵

Não é apenas a postura intencional que pode ser usada para previsão e explicação do ambiente circundante. Dennett propõe três tipos de atitudes perante algo que variam conforme o grau de precisão e de economia que buscamos. Quanto maior a precisão, menor a economia. As posturas menos precisas são úteis em diversas situações. Em alguns casos, a perda em termos de economia é tão grande que a busca da precisão deixa de valer a pena.

A primeira é a postura física. Ao adotar essa postura, explicamos e prevemos o comportamento de uma entidade a partir da análise da constituição física dessa entidade e das leis da física em geral. É uma maneira altamente precisa e segura de predição. A postura física poderia ser utilizada para predizer o comportamento de qualquer entidade, viva ou não. Porém, por ser um método laborioso, não é a postura mais indicada para prever o funcionamento de entidades vivas, por exemplo. Seria inútil tentar expressar com base em uma postura tão precisa como a postura física o comportamento de um sujeito consciente: “tão inútil como ler poemas em um livro por meio de um microscópio” (DENNETT, 1996/1997, p. 44).

⁸⁵ Essa parte sobre as três posturas propostas por Dennett apoia-se em trabalho realizado anteriormente (FAGUNDES, 2009).

Por exemplo, no caso de um despertador, podemos prever o seu comportamento com base em sua constituição física. Em geral, os relojoeiros fazem isso, mas não é uma postura adequada para os usuários comuns, pois demandaria muito tempo e estudo. É uma postura precisa, porém, pouco econômica. O gasto de tempo e de energia é muito alto. Contudo, para prever o comportamento de objetos que não são vivos e nem são artefatos humanos, uma pedra, por exemplo, a postura física é a única postura possível:

Quando predigo que uma pedra solta da minha mão cairá no solo, estou empregando a postura física. Não atribuo crenças e desejos à pedra; atribuo-lhe massa, ou peso, e me apoio na lei da gravidade para gerar minha predição. Para coisas que não são vivas e nem são artefatos, a postura física é a única estratégia disponível, embora possa ser empregada com vários níveis de detalhes, do subatômico ao astronômico (DENNETT, 1996/1997, p. 32).

A segunda postura que pode ser adotada é a postura de projeto [“design”]. Simplesmente, supomos que o objeto tenha um projeto específico, tendo em vista uma função, por exemplo, despertar no horário programado. As predições feitas com base em uma postura de projeto são mais arriscadas, pois pressupõem hipóteses como a de que uma entidade tenha sido projetada e que funcionará conforme esse projeto. No caso do despertador, podemos perder um compromisso importante se ele não funcionar conforme o projeto. Quando a postura física é adotada, não ocorrem erros de funcionamento. O risco da postura de projeto, contudo, é compensado pela facilidade e economia de tempo e energia ao fazermos a previsão.

Essa postura, de acordo com Dennett, pode ser adotada também em relação aos organismos vivos. Por exemplo, podemos olhar para uma árvore frutífera como um projeto da Mãe Natureza⁸⁶. A partir disso, prevemos que, se a regarmos apropriadamente, ela crescerá e poderá dar frutos, pois foi arquitetada dessa maneira. Para isso, não precisamos de informações detalhadas em física e química acerca do desenvolvimento das plantas.

Convém notar que muitas vezes a postura de projeto tem sido usada como um argumento teleológico para a existência de Deus. Há o famoso argumento do relógio, proposto por Paley em 1967 (ver: HIMMA, 2017). De acordo com o argumento, ao encontrar

⁸⁶ Para Dennett, um dos aspectos mais interessantes da teoria de Darwin é que ela explica o advento de projetos sem projetistas. “Mãe Natureza” é o próprio processo de seleção natural, sem uma inteligência adjacente a conduzi-lo. Vale a pena aqui apresentar uma citação de Dennett acerca de Spinoza: “Benedito Spinoza, no século XVII, associou Deus e natureza, afirmando que a pesquisa científica era o verdadeiro caminho da teologia. Por essa heresia ele foi perseguido. Existe uma perturbadora (ou, para alguns, atraente) dubiedade em sua visão herética de *Deus sive Natura* (Deus, ou Natureza): ao propor sua simplificação científica, ele estava personificando a natureza ou despersonificando Deus? A visão mais generativa de Darwin fornece uma estrutura em que podemos ver a inteligência da Mãe Natureza (ou essa inteligência é apenas aparente?) como uma característica não milagrosa e não-misteriosa – e portanto ainda mais maravilhosa – dessa coisa autocriadora” (1995/1998a, p.193, grifos do original).

um relógio no chão, uma pessoa que nunca tivesse visto um relógio na vida tenderia a atribuir àquele objeto uma função e um projeto feito por uma mente inteligente. O argumento se estenderia para o caso no qual a natureza é observada. Diante de sua diversidade e aparente harmonia, tenderíamos a lhe atribuir um projeto. Isso tem sido usado para afirmar a existência de um projetista anterior a toda a organização da vida. Assim, um aspecto importante da argumentação de Dennett é mostrar que a adoção da postura de projeto não implica na existência de um projetista⁸⁷.

Adotar a postura de projeto é muito útil para as nossas atividades quotidianas. Apesar de ser arriscada, esse tipo de postura governa as nossas relações com os artefatos e, em geral, estamos dispostos a arriscar muito por ela. Por exemplo, ao ligarmos um chuveiro elétrico, arriscamos até mesmo nossas vidas por confiarmos no projeto do chuveiro. Mesmo assim, a postura de projeto pode ser laboriosa demais para a previsão do comportamento de certos agentes. Deve-se notar que qualquer objeto explicável pela postura de projeto também pode ser explicado pela postura física, mas a recíproca não vale.

3.1.2. Sistemas intencionais

Dennett diz que o tipo mais simples de intencionalidade pode ser atribuído à chave e sua fechadura. A fechadura é capaz de discriminar a chave que a abre⁸⁸. Essa intencionalidade mais simples foi o elemento básico a partir do qual a natureza colocou os seus projetos em prática. Da mesma forma, os organismos unicelulares conseguem, na maior parte das vezes, discriminar alimentos e toxinas por um processo mecânico e simples. A partir do tipo mais simples de intencionalidade, conforme Dennett, surgiram seres com intencionalidade cada vez mais complexa. Por exemplo, algumas células se tornaram capazes de reconhecer a luz e se dirigir a ela no momento em que se foram equipadas com elementos fotossensíveis. Há organismos mais complexos que já nascem com a capacidade de representar a sua mãe, por exemplo. Alguns animais adquiriram capacidade de se camuflar, de modo a confundir outros sistemas intencionais.

⁸⁷ Em uma referência ao argumento de Paley, Dawkins (1986/2001) intitula seu livro de “O Relojoeiro Cego”. A evolução é uma designer que não consegue ver aquilo que está projetando, trabalha cegamente, mas produz funcionalidades.

⁸⁸ É preciso certo cuidado com esse exemplo de Dennett, pois a chave e a fechadura são feitas por um chaveiro inteligente para se direcionarem uma à outra. Um exemplo melhor é o da enzima e a molécula substrato sobre a qual ela age, desenhadas uma para a outra sem que ninguém as tenha desenhado. As enzimas possuem um formato complementar ao seu substrato, sendo esse complexo entre enzima e substrato geralmente compreendido como análogo ao complexo chave e fechadura.

Há ordens de intencionalidade. Conforme Dennett (1987), um sistema intencional de primeira ordem possui estados como crenças e desejos, mas não sobre outras crenças e desejos. Isto é, os sistemas intencionais de primeira ordem não possuem intencionalidade sobre a intencionalidade. Os sistemas intencionais de segunda ordem, por sua vez, possuem intencionalidade sobre sistemas intencionais: eles creem nas crenças de alguém, por exemplo. A intencionalidade de terceira ordem é a intencionalidade sobre a intencionalidade sobre a intencionalidade; por exemplo, quando João crê que Maria deseje que Cláudia goste de bolo; e assim sucessivamente. Provavelmente, há um limite na quantidade de ordens de intencionalidade com as quais nossa cognição é capaz de lidar, mas podemos imaginar uma enorme quantidade de ordens de intencionalidade.

As diversas habilidades cognitivas exigem diferentes ordens de intencionalidade. Conforme Abrantes (2013, p. 77), “agentes são entendidos como sistemas cognitivos cujo comportamento é causado por estados mentais”. Nesse sentido, os agentes possuem ao menos intencionalidade de primeira ordem. Já os intérpretes precisariam de intencionalidade de segunda ordem, uma vez que eles precisam atribuir intencionalidade a alguém.

Para explicar a ordem de intencionalidade necessária para haver comunicação, seja entre humanos ou entre animais não humanos, Dennett recorre a Grice. A existência da comunicação exige que um indivíduo x (aquele que está comunicando) queira que um indivíduo y (o que está ouvindo) reconheça que x quer que y produza uma resposta. Ou seja, a existência da comunicação exige que haja pelo menos intencionalidade de terceira ordem. Convém notar aqui que Dennett não está tratando especificamente da comunicação linguística, pois há seres não linguísticos que são capazes de se comunicar. No caso dos seres humanos, animais linguísticos, somos capazes de uma grande sobreposição de ordens de intencionalidade. No caso da comunicação linguística, torna-se necessário que um sujeito S (1) pretenda que o outro A (2) reconheça que S (3) pretende que A (4) creia que p. Então, para que a comunicação linguística possa ocorrer, Grice determina que são necessárias quatro ordens de intencionalidade, e Dennett (2017b) está de acordo.

3.1.3. A postura intencional

A terceira alternativa para a explicação e predição comportamental é, para Dennett, a adoção da postura intencional, a mais econômica e mais arriscada de todas elas, mas a que melhor funciona em muitos casos. Esse é o motivo pelo qual Dennett defende a sua adoção. A postura intencional pressupõe intencionalidade de segunda ordem, de modo a prever o

comportamento de uma entidade tratando-a como um agente racional. Isso não significa que a entidade seja um agente racional de fato. Ao adotar a postura intencional, essa questão não é colocada. Contudo, sem a adoção da postura intencional, é difícil prever o comportamento de certas entidades, em específico, dos sistemas intencionais.

Ao andar sobre um terreno acidentado, por exemplo, ajustamos a amplitude de nossos passos sem que precisemos pensar sobre isso (1996/1997, p. 19). De acordo com Dennett, uma onça calcula a trajetória do sol ao se deitar para um banho ou ao procurar a sombra para se disfarçar enquanto espreita sua presa, mas o seu cálculo não necessita de informações astronômicas detalhadas. A onça não pode falar acerca dos seus estados mentais, nem sobre como toma suas decisões comportamentais. Porém, podemos nos beneficiar da adoção de uma postura intencional se pretendermos explicar ou prever o seu comportamento. Podemos dizer, assim, que a onça deseja se disfarçar para que sua presa não a veja a tempo de fugir.

Um sistema cujo comportamento pode ser previsto por meio da adoção da postura intencional, ainda que não tenha consciência ou reflexão complexa acerca da intencionalidade que lhe é atribuída, pode ser considerado um sistema intencional conforme Dennett, desde que a intencionalidade possa explicar seu comportamento. Ou seja, intencionalidade e consciência não são a mesma coisa. A noção de consciência é considerada por Dennett excessivamente complicada. Um dos caminhos para descomplicar essa noção é mostrar que a intencionalidade não depende da existência de nenhuma outra coisa, ela se constrói passo a passo, existe em diversos graus e não é algo específico das mentes humanas.

A postura intencional, portanto, é uma proposta metodológica útil para explicarmos o funcionamento de certas entidades. Saber se aquela entidade possui intencionalidade ou não é o menos importante para Dennett. A postura intencional é importante porque ela funciona. Então, segundo o autor, a intencionalidade não é uma entidade com existência isolada dentro de um sujeito, mas sim algo que se atribui a diversas entidades a partir de uma perspectiva de quem pretende explicar e prever o seu comportamento. Um adulto pode explicar a uma criança o funcionamento de um despertador usando a postura intencional. Pode dizer, por exemplo, que nós usamos os botões para dizer ao despertador que horas são e a que horas queremos acordar. Então, como o despertador quer nos servir, ele tocará no horário solicitado.

No caso do despertador, portanto, vemos que qualquer das três posturas apresentadas pode ser adotada. Seu funcionamento pode ser explicado pela postura intencional, pela postura de projeto ou pela postura física. Porém, o uso da postura intencional para explicá-lo será extremamente limitado e superficial, já que o despertador não é um sistema intencional. Note-se, aqui, que as entidades explicáveis por meio da postura intencional também seriam

explicadas pela postura física ou pela postura de projeto, mas com ganho em termos de precisão na previsão e rigor explicativo - essa é uma possibilidade teórica, mas quase impossível do ponto de vista prático. Porém, nem todas as entidades que podem se sujeitar à postura de projeto ou à postura física poderão se sujeitar à postura intencional.

Assim, se estivermos falando do estudo da mente e do comportamento, contudo, é possível adotarmos uma postura física? Sim, é possível em princípio, mas é uma tarefa tão árdua, tão detalhada e complexa que se torna impraticável. O tempo de vida de um ser humano talvez não seja suficiente para explicar e prever uma única decisão de outro ser humano por meio da postura física. Nesse caso, seria necessário a reunião de um grande grupo de cientistas, cada um com conhecimentos parciais da situação⁸⁹, mas a postura intencional permite essa explicação e previsão de modo instantâneo e automático por qualquer um de nós.

Por isso, tentar explicar o comportamento humano pelos diversos aspectos físicos ou químicos envolvidos na situação seria um projeto frustrante. É despropositado ler um livro com um microscópio. Isso não significa atribuir características imateriais e irreduzíveis ao livro, mas sim adotar o distanciamento necessário à sua leitura. A adoção da postura intencional para o estudo da psicologia humana pressupõe um distanciamento em relação àquele corpo físico. Olhar para o mundo físico e químico para encontrar a mente humana é uma tarefa grandiosa demais e não vai nos levar a uma explicação coerente.

Então, embora a adoção da postura intencional não seja garantia da existência de intencionalidade, há determinados seres cujo comportamento só pode ser apropriadamente explicado e previsto por meio da adoção dessa postura, são os sistemas intencionais. O sucesso na atribuição de crenças e desejos quando as demais posturas se tornam inúteis é, assim, o instrumento apropriado para a detecção de estados mentais em um sistema.

3.1.4. Intencionalidade intrínseca e derivada

Foi dito⁹⁰ que intencionalidade é a característica dos estados mentais (ou pelo menos de alguns deles) de estarem direcionados a algo. Em geral, é atribuída apenas aos estados

⁸⁹ Dennett (2017b, p. 375-376, tradução nossa) menciona o fato de que o conhecimento, da maneira como é produzido atualmente nas ciências, muitas vezes não cabe em um único sujeito. São cientistas de diversas áreas que trabalham coletivamente para produzir um trabalho de múltiplos autores no qual “muitos dos detalhes são apenas parcialmente compreendidos por cada autor”. No caso de uma compreensão científica mais ampla da consciência, segundo ele, seria necessário um trabalho desse tipo. É uma compreensão que envolve muitos detalhes e um sujeito apenas não seria capaz de lidar cognitivamente com tanta informação. Porém, vale notar que nossa compreensão da consciência por meio da postura intencional tem funcionado muito bem para as situações enfrentadas quotidianamente.

⁹⁰ Ver: seção 1.5.

mentais. Mas o que dizer de um livro ou de uma fotografia, por exemplo? É notório que um livro é sobre algo, assim como uma fotografia é uma fotografia de algo. Contudo, um livro ou uma fotografia não parecem possuir mente. Nesse sentido, eles possuem intencionalidade? Alguns autores, em especial John Searle (2002), argumentam que há uma diferença entre intencionalidade intrínseca e derivada. A intencionalidade do livro, por exemplo, seria derivada das intencionalidades intrínsecas do autor e do leitor. Searle diferencia também intencionalidade literal e metafórica. Segundo ele, em uma crítica dirigida a Dennett, ao atribuir intencionalidade a um termostato, por exemplo, estaríamos diante de um caso de intencionalidade metafórica.

Tanto as crenças quanto as experiências visuais são fenômenos intencionais intrínsecos nas mentes/cérebros dos agentes. Dizer que elas são intrínsecas é apenas dizer que os estados e eventos realmente existem nas mentes/cérebros dos agentes; as descrições desses estados e eventos devem ser tomadas literalmente, e não apenas como um modo de falar, nem como uma forma abreviada de uma descrição de um conjunto mais complexo de eventos e relações ocorrendo fora do agente (SEARLE, 2002, p.78, tradução nossa).

De acordo com a teoria dos sistemas intencionais de Dennett, por outro lado, não há distinção teoricamente motivada que permita traçar um limite entre intencionalidade intrínseca e derivada nem entre os casos literais e os metafóricos. No exemplo da atribuição de uma preferência pelo sol ao manjericão, aparentemente estamos diante de um caso em que a intencionalidade não é intrínseca, já quando atribuímos intencionalidade a um interlocutor, tudo parece estar perfeitamente no seu lugar, assim como tudo parece certo quando atribuímos intencionalidade ao macaco que rouba as bolsas dos visitantes em um parque⁹¹, mas o que dizer de um robô jogador de “poker” que é capaz de blefar?⁹²

Para quem defende a distinção entre intencionalidade original e derivada, a original já estaria na mente do sujeito, e a intencionalidade atribuída, portanto, seria derivada. Isso nos leva ao ceticismo, pois apenas o sujeito poderia ter intencionalidade original, já que no

⁹¹ No Parque Nacional Água Mineral, em Brasília, os macacos às vezes roubam as bolsas dos visitantes em busca de alimentos. Em algumas ocasiões, eles se aproximam em grupo de alguns visitantes que tenham se sentado desavisadamente sob uma árvore e começam a gritar. Diante desse comportamento do macaco, o visitante tende a se assustar e sair de perto, deixando a bolsa e os alimentos disponíveis (exceto os visitantes experientes, que levam suas bolsas junto ao saírem de perto dos macacos).

⁹² Em tese, é ainda mais difícil para um computador ganhar no “poker” do que no xadrez, por duas razões: primeiro porque, no “poker”, as cartas ficam escondidas nas mãos dos jogadores, enquanto no xadrez é possível ver as posições das peças no tabuleiro e as jogadas executadas pelo adversário, o que permite a qualquer jogador (humano ou não) calcular as possibilidades de movimentos e probabilidades de sucesso; segundo porque não é possível obter uma boa taxa de sucesso no “poker” sem blefar. Porém, sabe-se que há robôs competindo com as pessoas nos jogos “on line”, os chamados “pokerbots”, passando-se por jogadores humanos. Esses programas jogadores de poker estão disponíveis na internet, sendo usados por trapaceiros e, quando descobertos, são banidos dos sites de jogos. Também há jogadores humanos que utilizam os “bots” para estudos, com o intuito de melhorar sua própria estratégia ao enfrentarem adversários humanos.

processo de comunicação com outros, nós lhes atribuímos intencionalidade. Se formos levar a distinção entre intencionalidade original e derivada às últimas consequências, a intencionalidade original ficaria restrita apenas ao sujeito ensimesmado.

Se a atribuição de estados mentais é um recurso de interação tanto para nós quanto para outros animais, podemos supor que, na história de vida (ontogenia) de cada animal capaz de atribuir estados mentais, essa atribuição esteja na base de suas relações com o mundo. A auto-atribuição de intencionalidade seria, então, algo posterior e derivado, por sua vez, da atribuição de intencionalidade aos outros. Por isso, nada há de absurdo em imaginar que a atribuição de intencionalidade venha antes da auto-atribuição. Nesse caso, então, a intencionalidade original seria a dos outros e a derivada seria a do próprio sujeito? Traçar uma linha entre intencionalidade original e derivada não é uma tarefa nada simples. Se adotamos uma perspectiva como a de Dennett, não faz sentido tentar traçar essa linha, pois a intencionalidade é algo que se atribui aos agentes com os quais interagimos, não é algo original nem intrínseco.

Na concepção de Dennett, nossa intencionalidade resulta de um processo pelo qual ancestrais com sistemas sensoriais mais simples começam a discriminar o mundo. Não é possível traçar uma linha demarcatória do momento a partir do qual a intencionalidade começa a surgir na história dos seres vivos, é um processo gradual. Por exemplo, imagine uma planária se movendo em um recipiente em direção à área mais rica em nutrientes. O comportamento dela, conforme o autor, é explicável conforme o seguinte *princípio de racionalidade*: ela está procurando o próprio bem. Se essa procura é literal ou metafórica, é algo que não se pode determinar de modo não arbitrário. Para Dennett, o sucesso na adoção da postura intencional pode ser considerado um excelente indício de intencionalidade real, não fazendo sentido diferenciar a intencionalidade original e a derivada. O termo “real”, aqui, não deve ser tomado no sentido de que haja algo lá dentro do agente que se possa encontrar e que seja a intencionalidade daquele agente, mas sim que a intencionalidade é, de fato, algo que se atribui aos sistemas intencionais e, considerando o sucesso obtido nessa atribuição, eles *são* intencionais.

Miguens (2002) faz uma crítica a Dennett em relação a isso. Ela defende que a distinção entre intencionalidade intrínseca e derivada não pode ser totalmente abandonada por quem pretende adotar uma perspectiva interpretivista. Segundo ela, se essa distinção for abandonada, a noção de intérprete se torna confusa, pois é o intérprete quem atribui estados mentais. Porém, é importante notar que, numa perspectiva dennettiana, os estados mentais se sustentam primeiramente na medida em que são atribuídos a um sistema intencional. O

intérprete, por ser um sistema intencional, também é sujeito à atribuição de estados, inclusive por si mesmo. Não há, no pensamento de Dennett, uma primeira e originária fonte de intencionalidade. Ela surge gradualmente ao longo do processo evolutivo. Quem atribui intencionalidade também está sujeito à atribuição de intencionalidade. Por isso, o intérprete não pode ser a fonte original de toda intencionalidade⁹³.

3.1.5. *A postura intencional e o princípio de caridade*

É possível observar importantes semelhanças entre a postura intencional de Dennett e o princípio de caridade de Davidson⁹⁴. O princípio de caridade consiste na atribuição de racionalidade e coerência ao interlocutor, tal qual a postura intencional. Porém, a postura intencional é aplicada de modo bastante amplo, para compreender e prever o comportamento de muitos agentes. Davidson, por sua vez, está preocupado apenas com a explicação e previsão do comportamento humano, o que gera uma lacuna na compreensão de como a evolução pode passar do não pensamento para o pensamento.

Davidson (1975/1984e), inclusive, compreende a atividade de pensar como algo que depende da capacidade de interpretação linguística. A interpretação linguística, por sua vez, só pode ser compreendida de modo relacional com uma rede de crenças de fundo e um conjunto de crenças compartilhadas as quais, mesmo que o intérprete não esteja pensando sobre elas, servem de cenário para que a interpretação seja possível. Essas crenças envolvem distinções muito finas e abstratas que, na concepção de Davidson, os animais não são capazes de fazer. Embora sejam bastante esclarecedoras acerca da linguagem e do pensamento proposicional, as pesquisas de Davidson se restringem à mente humana⁹⁵. Ele compreende o pensamento como proposicional e foca seus estudos nesse aspecto do mental.

Pode-se acreditar que Scott não seja o autor de *Waverley* sem duvidar que Scott seja Scott; pode-se querer ser o descobridor de uma criatura com um coração sem que se queira descobrir uma criatura com um rim. Pode-se pretender morder a maçã na mão sem que se pretenda morder a única maçã que tem um bicho nela, e por aí vai. A

⁹³ A perspectiva de Dennett acerca do mental pode ser considerada externista, assim como a de Davidson. Foi dito que, de acordo com o externismo, ao menos alguns tipos de estados mentais, como crenças, por exemplo, dependem de uma determinada relação com o ambiente externo (LAU; DEUTSCH, 2016). Considerando isso, vale notar que, seguindo o pensamento externista, a distinção entre intencionalidade intrínseca e derivada provavelmente perde sua razão de ser já que muitos dos estados mentais são definidos a partir de uma relação causal entre a representação e aquilo que é representado, portanto não são intrínsecos. Essa perda do sentido da distinção depende, convém observar, da adoção prévia desse tipo de abordagem dos estados mentais. A refutação de Putnam (1981) ao argumento do cérebro numa cuba depende disso.

⁹⁴ É curioso que Davidson não seja mencionado por Dennett no livro “Ideia Perigosa de Darwin” (1995/1998a), trabalho em que a postura intencional é bastante desenvolvida. A influência de Davidson parece ser presente ali.

⁹⁵ Ver seção 2.3.

intencionalidade⁹⁶ que tanto fazemos na atribuição de pensamentos é muito difícil de se fazer quando o discurso não está presente. O cachorro, nós dizemos, sabe que seu mestre está em casa. Mas ele sabe que o sr. Smith (que é o seu mestre), ou que o presidente do banco (que é esse mesmo mestre) está em casa? Não temos ideia de como resolver, ou dar sentido, a essas questões (DAVIDSON, 1975/1984e, p. 163, tradução nossa).

Quanto a Dennett, a postura intencional não compromete o investigador com uma hipótese específica acerca de estruturas internas subjacentes às competências daquele agente. Ela é neutra em relação a isso, o agente não precisa ser humano nem ter uma estrutura física orgânica. É preciso, sim, haver uma estrutura física suficientemente complexa para instanciar o tipo e a ordem de intencionalidade daquele organismo, mas o que realmente importa para que se identifique um sistema intencional é que suas ações possam ser previstas por meio da atribuição de crenças e desejos. Vale notar que um aspecto enfatizado por Dennett acerca da estrutura de um agente intencional é que ela seja *fisicamente possível*. Experimentos de pensamento que incluem situações logicamente possíveis e fisicamente impossíveis, na concepção de Dennett, não servem como objeções à teoria dos sistemas intencionais, uma vez que a teoria se apoia nesse postulado da possibilidade física.

Embora a postura intencional não tenha nada a dizer sobre o que está acontecendo dentro do agente, ela esclarece os mecanismos cognitivos daquele agente. Nesse sentido, vai além da mera atribuição. Atribuímos, sim, intencionalidade a entidades que não são intencionais, mas certas entidades apresentam um comportamento que não faz sentido senão por meio dessa atribuição, essas entidades possuem competências cognitivas que dependem de que tenham intencionalidade.

A postura intencional é, portanto, um modo teoreticamente neutro [“theory-neutral”] de *capturar as competências cognitivas* de diferentes organismos (ou outros agentes) sem comprometer o investigador com hipóteses excessivamente específicas sobre as estruturas internas que subjazem às competências (DENNETT, 2009, p. 344, grifos nossos, tradução nossa).

Observe-se que, em muitas das vezes, por mais obviamente racional que seja o agente cujo comportamento está sendo explicado, aquelas atitudes proposicionais que lhe são atribuídas não são conscientes. Por exemplo, posso dizer que um motorista jogou o carro no acostamento porque quis evitar a colisão com um caminhão que vinha pela contramão na curva e a pessoa só viu o caminhão quando ele estava muito próximo. Porém, embora essa atribuição de uma série de deliberações ao motorista seja perfeitamente legítima, as deliberações foram feitas de modo tácito, sem que elas estivessem presentes à sua consciência.

⁹⁶ Ver: nota de rodapé 35, Capítulo 1.

“A postura intencional é maximamente neutra sobre como (ou onde, ou quando) o trabalho duro da cognição é terminado, mas garante que o trabalho *é* feito pelo teste de sucesso” (DENNETT, 2009, p. 348, grifo do original, tradução nossa).

3.1.6. *Intencionalidade da natureza*

A teoria dos sistemas intencionais, segundo Dennett (2009), embora não seja útil para explicar os mecanismos internos dos sistemas intencionais, serve para explicar como e porque somos capazes de dar sentido ao comportamento dos sistemas complexos. Além disso, ela preenche aparentes descontinuidades entre as mentes animais mais simples e as mentes humanas. Pode ser usada em diversas disciplinas: na ciência da computação, na biologia evolutiva, na psicologia animal, mas também tem uso indispensável em nossas relações cotidianas com outros sistemas intencionais.

Dennett nota que a evolução dos seres vivos pode ser explicada com sucesso por meio da adoção da postura intencional. Para ele, a evolução tem intencionalidade sem que tenha propósito nem inteligência. A própria evolução é direcionada para a adaptação, mas esse direcionamento é cego. Dawkins (1996/1998) apresenta a evolução como um trajeto direcionado à escalada de um monte. Essa ideia é uma resposta ao argumento de que a harmonia e diversidade dos seres vivos não poderia ser desenhada sem que houvesse uma inteligência divina conduzindo o processo. O monte é improvável porque é escalado a partir da própria ordem natural, sem um guia inteligente conduzindo o processo de fora.

[...] Darwin sugere uma divisão: dê-me Ordem, ele disse, e tempo, e eu lhe darei o Projeto. Deixe-me começar com uma regularidade – a regularidade sem propósito, irracional e sem objetivo da física – e eu lhe mostrarei um processo que acabará produzindo coisas que exibem não só intencionalidade como um projeto intencional. (Foi exatamente isso que Karl Marx pensou ter visto ao afirmar que Darwin havia dado um golpe mortal na teleologia: Darwin *reduzira* teleologia a não-teleologia, Projeto a Ordem.) (DENNETT, 1995/1998a, p. 68, grifos e maiúsculas do original).

O que Dennett procura defender é que a natureza é uma artesã sem discernimento, mas extremamente hábil e muito mais engenhosa do que nós, os supostos seres dotados de uma intencionalidade original. Porém, no fim das contas, nossa intencionalidade também é o resultado de uma série de mecanismos simples. O comportamento de cada um desses mecanismos, se tomado isoladamente, poderia ser explicado simplesmente pela adoção da

postura física⁹⁷. Porém, a união desses mecanismos de maneira complexa dá origem à intencionalidade.

Repetidas vezes, biólogos desconcertados diante de alguma partícula aparentemente fútil ou desastrosa de um projeto ruim na natureza acabaram vendo que haviam subestimado a engenhosidade, o brilho puro, a profundidade de visão a ser descoberta em uma das criações da mãe natureza. Francis Crick maliciosamente batizou essa tendência com o nome de seu colega Leslie Orgel, falando do que ele chama de “segunda lei de Orgel: a evolução é mais esperta que você” (DENNETT, 1995/1998a, p. 77, aspas do original).

Dennett defende que a intencionalidade não é algo que surge repentinamente na história evolutiva iluminando a mente humana que, por sua vez, passa a iluminar as outras coisas conferindo-lhes intencionalidade. Ao contrário, ela surge gradualmente, a partir de formas mais simples de intencionalidade, como a ameba que se direciona ao nutriente e se afasta do meio tóxico. A intencionalidade “se infiltra de baixo para cima, desde os processos algorítmicos inicialmente irracionais e inúteis que, ao se desenvolver, vão aos poucos adquirindo significado e inteligência” (DENNETT, 1995/1998a, p. 213)⁹⁸. Isso ocorre no desenrolar temporal da evolução, mas também ocorre a partir da complexidade espacial sincrônica de diversos subsistemas ignorantes agindo concomitantemente e fazendo funcionar os organismos mais complexos (como ocorre no modelo dos rascunhos múltiplos).

[...] em um organismo com genuína intencionalidade – como você mesmo – existem, neste exato instante, muitas partes, e algumas delas exibem uma espécie de semi-intencionalidade – ou mera *como se fosse* intencionalidade, ou pseudo-intencionalidade – chame-a como quiser – e a sua própria intencionalidade genuína, amadurecida, é de fato o produto (sem nenhum outro ingrediente milagroso) das atividades de todos os pedacinhos semi-rationais e irracionais de que você é composto. (DENNETT, 1995/1998a, p. 213, grifos do original.)

De acordo com Dennett, Darwin descreveu o processo evolutivo por meio de um processo algorítmico. Essa descrição é extremamente simples e plausível e o mesmo mecanismo, descrito de modo algorítmico, segundo Dennett, pode ser estendido a diversos outros domínios, incluindo a cosmologia e a psicologia. Dennett, (1995/1998a) considera a ideia de Darwin como um “ácido universal”, capaz de penetrar e transformar profundamente a nossa visão de mundo. É essa ideia que permitiu a existência dos projetos da natureza sem a necessidade de um projetista inteligente a conduzir o processo.

A ideia de Darwin nascera como uma resposta às perguntas da biologia, mas ameaçava vazar, oferecendo respostas - bem-vindas ou não - para dúvidas existentes na cosmologia (de um lado) e na psicologia (de outro). Se o *reprojeto* podia ser um processo irracional, algorítmico, de evolução, por que o processo todo não poderia

⁹⁷ Ver seção 3.1.1.

⁹⁸ Sobre o algoritmo evolutivo, ver seção 1.4.1.

ser o produto da evolução, e daí por diante, *até embaixo*? E se a evolução irracional pode ser responsável pelos artefatos surpreendentemente inteligentes da biosfera, como os produtos das nossas próprias mentes ‘reais’ poderiam estar isentos de uma explicação evolutiva? A ideia de Darwin, portanto, também ameaçava se espalhar *até em cima*, desfazendo a ilusão de nossa própria autoria, nossa própria centelha divina de criatividade e compreensão (DENNETT, 1995/1998a, p. 66, grifos do autor).

3.1.7. *Mentes dentro de mentes*

É possível explicar a ascensão gradual do tipo mais simples de intencionalidade até a intencionalidade das mentes humanas, por meio de um modelo que Dennett (1996/1997) denomina “torre de gerar e testar”⁹⁹, no qual os organismos mais simples começam a introjetar processos cognitivos de tipo darwinista em camadas. Isso confere ao processo evolutivo o surgimento de organismos com novas habilidades cognitivas para lidar com o ambiente, progressivamente mais complexas. Contudo, os processos cognitivos sempre seguiriam o mesmo algoritmo que está em funcionamento na evolução dos seres vivos. A sobreposição desse algoritmo em ambientes cognitivos internos permitiria arquiteturas cognitivas cada vez mais elaboradas, até que uma mente cultural e linguística, como a humana, tivesse condições de surgir.

No primeiro andar da torre se encontram as criaturas darwinianas. São simples replicadores que enfrentam diretamente o crivo do ambiente. Suas competências já estão completamente pré-fixadas, não possuem nenhuma flexibilidade comportamental. A qualquer momento, podem morrer sem conseguirem se replicar, mas também podem deixar várias réplicas igualmente rígidas. Há os erros de cópia que produzem variantes aleatoriamente, as quais são testadas diretamente pelo ambiente, com grande chance de reprovação no teste, implicando na completa extinção daquela variante. Essas criaturas se direcionam para o que lhes é favorável ao longo das gerações, de modo totalmente mecânico.

No segundo andar da torre estão as criaturas denominadas skinnerianas. Elas possuem disposições comportamentais inatas, assim como as darwinianas. Porém, são equipadas com mecanismo de reforço cognitivo que lhes permite certo grau de ajuste comportamental. Elas produzem reações comportamentais que, em um primeiro momento, são aleatórias, mas podem receber reforço positivo ou negativo em confronto com o ambiente, levando-as a uma variação nas taxas das respostas comportamentais apresentadas em seguida. Nesse sentido,

⁹⁹ A torre de gerar e testar também é apresentada em Fagundes (2009), com diferente ênfase. Aqui se pretende pensar sobre como a intencionalidade pode ter se infiltrado gradualmente ao longo do processo evolutivo.

seu comportamento acaba se direcionando a respostas mais adaptativas de uma forma extremamente simples. “Aqueles variantes nascidas com a infeliz disposição a confundir estímulos positivos e negativos, fugindo das coisas boas e indo para as coisas ruins, logo irão eliminar a si mesmas, não deixando prole” (DENNETT, 2017b, p. 98, tradução nossa). Embora essa capacidade das criaturas skinnerianas lhes confira uma maior flexibilidade comportamental e melhor adaptabilidade, é ainda bastante arriscada, pois a criatura precisará produzir o comportamento de modo aleatório e depois confrontá-lo com o ambiente.

Em seguida surgem as criaturas popperianas. Elas possuem uma representação interna do ambiente externo. Isso lhes permite testar as possibilidades comportamentais internamente, prevendo as reações que serão apresentadas pelo ambiente a cada uma de suas possíveis escolhas comportamentais. Suas escolhas diante do mundo real deixam assim de ser aleatórias. Elas passam a se direcionar cognitivamente àquilo que foi testado previamente.

No último andar da torre estão as criaturas gregorianas. São as criaturas que, além de possuírem as habilidades cognitivas das criaturas popperianas, vivem em um ambiente cultural, não apenas um ambiente natural. Essas criaturas foram equipadas com “instrumentos de pensamento”, sendo a linguagem o mais importante desses instrumentos. As criaturas gregorianas, na concepção de Dennett, possuem uma arquitetura cognitiva que permite o surgimento de um novo processo evolutivo: a evolução cultural. Esse processo possui suas próprias variantes, não mais as variantes biológicas do processo de seleção natural e possui um ambiente específico, o ambiente da linguagem e das instituições humanas. Porém, é considerado pelo autor um processo tipicamente darwinista por possuir abstratamente as três características que fazem rodar o algoritmo evolutivo¹⁰⁰. Apenas após o advento da cultura, segundo Dennett, é que podem surgir mentes como as humanas.

É importante notar que cada andar da torre só pode ser construído sobre uma base bem alicerçada. Por isso, as criaturas skinnerianas também são darwinianas, as criaturas popperianas também são skinnerianas e as criaturas gregorianas também são popperianas. Nós, portanto, reunimos as características de todas essas criaturas, mas somos os únicos capazes de implementar o processo de evolução cultural. Isso, para o autor, faz com que nossas mentes sejam completamente diferentes. Mais à frente, esse tema será retomado. A seção a seguir trata da atribuição de estados mentais com um pouco mais de profundidade.

¹⁰⁰ Dennett (2017b), após aprofundar os estudos sobre evolução cultural por meio de diversos trabalhos desenvolvidos desde a publicação de “Darwin’s Dangerous Idea” (1995/1998a), percebe que o algoritmo evolutivo é ainda um modelo muito simplista, tanto para a evolução cultural quanto para a biológica. Há muitos outros tipos de processos acontecendo, como teria apontado Godfrey-Smith (2007 *apud* DENNETT 2017b) com sua noção de espaços darwinianos. Porém, não entraremos aqui nesse debate.

3.2. Teoria da mente

No final da década de 70, os psicólogos Premack e Woodruff (1978) desenvolveram alguns testes para detectar se os chimpanzés possuíam a capacidade de atribuir estados mentais a outras criaturas, chamada por eles de “teoria da mente”. O macaco era exposto a um vídeo no qual um humano aparecia com um problema para resolver, por exemplo, um cacho de bananas muito alto para ser alcançado. Em seguida, eram oferecidas duas fotos ao macaco com possíveis soluções para aquele problema. Nesse exemplo, uma foto era do humano com um tamborete e a outra era do humano com um cabo de vassoura. O macaco deveria escolher uma das fotos. O tamborete, nesse caso específico, seria mais útil, pois permite ao humano subir e alcançar as bananas. (O cabo de vassoura seria mais útil se o objetivo fosse puxar as bananas de baixo de um móvel, por exemplo.) As escolhas efetuadas pelos chimpanzés sugeriam capacidade de atribuir ao personagem do vídeo o propósito de solucionar o problema e a capacidade de escolher alternativas apropriadas para o alcance desse propósito. Esse trabalho foi bastante influente no debate acerca da psicologia de senso comum e da atribuição de estados mentais.

Ter uma teoria da mente, segundo alguns autores, é poder atribuir estados mentais, como desejos, crenças e temores a outros seres, entendendo (1) que pode haver uma divergência entre esses estados mentais atribuídos e a realidade externa e (2) que os estados mentais atribuídos possuem relação causal com o comportamento daquele ser. Dennett, por sua vez, considera que fazemos isso de modo tão natural que a noção de “teoria” não caberia, pois uma “teoria” deveria poder ser formulada linguisticamente e as atribuições de estados mentais que fazemos são frequentemente muito automáticas e não proposicionais. Por isso, Dennett prefere falar em psicologia do senso comum.

Alguns pesquisadores gostam de chamar psicologia de senso comum de “teoria da mente” (ou apenas TOM), mas isso me parece enganoso, uma vez que isso tende a prejudicar a questão de como nós conseguimos ter um tal talento, sugerindo que nós temos, e usamos, uma teoria. Em um espírito similar, nós teríamos que dizer que você tem uma *teoria da bicicleta* se você sabe como andar de bicicleta, e é a sua *teoria da nutrição* que explicaria sua habilidade de evitar morrer de fome e evitar comer areia. Isso não me parece uma maneira útil de pensar sobre essas competências. Uma vez que todos concordam que nós temos o talento interpretativo, e todas as pessoas discordam sobre como nós conseguimos ser tão competentes, eu penso que é melhor manter o termo “teoria” fora disso e usar um termo um pouco mais neutro por enquanto. A psicologia acadêmica ou científica também estão no negócio de explicar e prever as mentes dos outros, e ela realmente tem teorias: behaviorismo, cognitivismo, modelos neurocomputacionais, psicologia Gestalt e uma série de outras. A psicologia de senso comum é um talento em que nós nos destacamos sem educação formal (DENNETT, 2013, p. 74-75, grifos e aspas do original, tradução nossa).

Outro teste que tem sido muito utilizado para verificar a existência de capacidade de atribuição de intencionalidade é o de falsa crença, desde que foi proposto por Wimmer e Perner (1983). No teste, o sujeito é colocado diante da situação em que um ator x coloca um objeto em uma posição e sai de cena. Sem que o ator x veja, entra o ator y e muda o objeto de posição. Pergunta-se ao sujeito submetido ao teste onde o ator x procurará o objeto quando retornar. As crianças de até 4 anos apontaram para o lugar onde o ator y deixou o objeto, ou seja, não puderam atribuir ao ator x , protagonista da cena, a falsa crença de que o objeto ainda estaria onde ele (x) o deixou. Porém, a partir dos quatro anos, algumas crianças começam a apontar corretamente para o local onde o ator y havia deixado o objeto. Com seis anos, a grande maioria das crianças já consegue passar no teste de falsa crença.

Até recentemente, acreditava-se que a atribuição de falsas crenças era uma habilidade especificamente humana, mas o teste foi aplicado por um grupo de pesquisadores (do qual Tomasello faz parte) em três espécies de símios (chimpanzés, bonobos e orangotangos), e os resultados sugerem que eles também podem atribuir falsas crenças. O teste era feito pelo mapeamento da direção do olhar dos macacos. “A compreensão da falsa crença tem interesse particular porque requer o reconhecimento de que as ações dos outros são dirigidas não pela realidade, mas pelas crenças acerca da realidade, mesmo quando essas crenças são falsas” (KRUPENYE; KANO et al., 2016, tradução nossa).

Na interpretação dos resultados, os autores demonstram não saberem se os macacos percebem o que estão fazendo com o olhar, consideram que os macacos demonstraram uma habilidade implícita de detecção de falsas crenças, menos flexível do que a desenvolvida por humanos. “Dado que os símios ainda não obtiveram sucesso em tarefas que meçam as falsas crenças com base em escolhas comportamentais específicas, a presente evidência pode constituir um entendimento implícito da crença” (KRUPENYE; KANO et al., p. 113, tradução nossa). Passar no teste de falsa crença revela que aquele sistema é capaz de atribuir intencionalidade e de separar estados do mundo externo dos estados daquele sujeito que está diante dele.

3.2.1. A Psicologia de Senso Comum

Desde a década de 80, a noção de psicologia do senso comum tem se mostrado bastante central nos debates em ciências cognitivas e filosofia da mente (RAVENSCROFT, 2016). Psicologia de senso comum é a nossa prática quotidiana de atribuição de estados intencionais aos outros para explicar e prever o seu comportamento. Para alguns, ela é

considerada uma teoria da mente, no sentido que a atribuição de crenças é feita em forma de uma crença: “Creio que meu amigo crê que hoje vai chover.” Ou seja, a psicologia de senso comum pode ser vista como uma teoria que alguns seres possuem acerca dos estados mentais alheios e serve para explicar e prever o seu comportamento. Dennett, contudo, vê a psicologia de senso comum como uma habilidade pré-linguística. Por essa razão, não a compreende como uma teoria, sim algo que fazemos de modo natural e automático.

O termo “psicologia de senso comum” (“Folk Psychology”) é usado por Dennett desde 1981 para designar o que todas as pessoas sabem acerca das mentes, das próprias, mas principalmente das mentes alheias. Dada a pouca informação que temos do que se passa nas mentes dos outros, é impressionante o grau de confiança que habitualmente depositamos em nossas atribuições de estados mentais. Note-se que a psicologia de senso comum não precisa restringir a atribuição de estados mentais apenas aos seres humanos. A teoria dos sistemas intencionais de Dennett considera a psicologia de senso comum como uma habilidade que se manifesta pela adoção de determinada postura diante de outros seres: a postura intencional.

Um ponto a ser destacado aqui acerca da teoria dos sistemas intencionais de Dennett é que, para ele, ter crenças não é algo que se defina pela capacidade de atribuir crenças a alguém. Essa é uma crítica que ele faz tanto a Davidson quanto a Premack. Na concepção de Dennett, esses dois autores erram por perseguirem a essência das crenças, ou as condições necessárias e suficientes para que alguém possa ser considerado um verdadeiro crente [“true believer”]. Davidson, nesse sentido, restringe a capacidade de crer aos seres humanos. A concepção de crença de Dennett é bastante deflacionada quando comparada com essa que pressupõe a capacidade de atribuição de crenças. No fim das contas, a crença é, para ele, um estado cognitivo capaz de influenciar o comportamento de um sistema, ainda que não seja um estado proposicional.

Ocorre que a psicologia de senso comum está na base de todas as nossas relações humanas e de quase todos os nossos projetos. Ela provavelmente se desenvolve na linhagem hominídea como uma resposta à crescente complexidade do ambiente social de nossos antepassados (ABRANTES, 2006). Além disso, o seu poder preditivo é extremamente poderoso, servindo para maior parte de nossas ações quotidianas. Mesmo assim, conforme Abrantes (2013, p. 84):

[h]á uma antiga e persistente controvérsia na literatura em etologia cognitiva, assim como nas resenhas filosóficas dessa literatura, a respeito de ser ou não adequada uma linguagem intencional (tomada, basicamente, da psicologia de senso comum) para descrever e, possivelmente, explicar o comportamento de animais não humanos.

O risco, nesse caso, seria cair em um antropocentrismo ou, na melhor das hipóteses, em um “primatocentrismo”. Dennett, com sua postura intencional, defende a atribuição de atitudes proposicionais às mais diversas entidades, dentre as quais, algumas não são sequer orgânicas. Ele apresenta a psicologia de senso comum não como uma teoria, mas como uma arte ou uma habilidade de fazer algo [“craft”]. Com isso, ele quer dizer que não estamos diante de um conhecimento proposicional, mas sim de um saber fazer, algo que se aprende praticando (DENNETT, 1998b, p. 83 e segs.).

Quando falamos sobre a teoria da mente, para o autor, nossa fala vem repleta de preconceitos, chamados por ele de ideologia de senso comum, os quais não se refletem na prática quotidiana de atribuição de estados mentais. Curiosamente, portanto, a arte da atribuição de atitudes proposicionais não seria para ele um conhecimento proposicional, mas uma arte dominada por todos nós, uma espécie de sabedoria de vida que aprendemos nas relações com outras pessoas, inclusive relações que não são linguísticas. “O que nós aprendemos em idade tenra [“at mother’s knee”] enquanto estamos crescendo e que poderia ser em algum grau inato é um talento variado para ter expectativas sobre o mundo. Muito disso nunca se torna articulado em nada remotamente parecido com proposições” (DENNETT, 1998b, p. 82, tradução nossa). A psicologia de senso comum não pode ser uma teoria, segundo Dennett, pois não é formulada em termos de leis ou teoremas explícitos. Ela é uma habilidade.

Como é que “todos sabem” a psicologia de senso comum? Ela é, no fim das contas, ToM¹⁰¹, um tipo de teoria que você aprende quando criança? Algo dela é inato, ou se a aprendemos toda, como e quando o fazemos? Houve um fluxo de pesquisas sobre essas questões nos últimos trinta anos. Minha resposta (ainda, depois de todos esses anos de pesquisa), é que ela não é tanto uma teoria como é uma prática, uma forma de investigar o mundo que vem tão naturalmente que deve ter alguma base genética em nossos cérebros. Você tem que aprender algo dela em idade tenra [“at mother’s knee”], e se você fosse de algum modo privado de todo contato com outros seres humanos enquanto crescesse, você provavelmente seria surpreendentemente inepto em psicologia de senso comum (junto com suas outras deficiências sérias), mas o impulso é muito forte de interpretar coisas que se movem de modo irregular (diferentemente de um pêndulo ou uma bola rolando por uma colina) como agentes (DENNETT, 2013a, p. 75, aspas do original, tradução nossa).

Por que isso faz sentido? Primeiro porque, em geral, nós não transformamos nossas atribuições de estados mentais aos outros em proposições. Ao contrário, fazemos isso automaticamente o tempo todo e tomamos decisões comportamentais a partir daí. Segundo que não atribuímos às pessoas crenças e desejos discretos. Para que possamos pensar uma proposição como “meu amigo pensa que eu gosto de lasanha”, é preciso atribuir ao mesmo

¹⁰¹ ToM: “Theory of Mind”, ou seja, teoria da mente.

tempo a ele uma quantidade enorme de outras crenças, como “ele tem as mesmas crenças que eu acerca do que a maior parte das palavras significam”, “ele deseja me atribuir um gosto por algo”, “ele crê que lasanha seja algo bom para se comer, ou que seja ao menos comestível” etc. Além disso, provavelmente também seja necessário atribuir a ele estados mentais e habilidades cognitivas não proposicionais (desde que disposicionais, claro! *Qualia* não relacionais estão fora!). Mas esse conjunto de atitudes proposicionais e demais estados mentais não podem ser atribuídas de modo discreto, uma a uma, ou teríamos um problema de “explosão combinatória” (DENNETT, 1998b, p. 93).

Um exemplo bastante ilustrativo dado por Dennett (2013c) é o dos preconceitos. Ele pede para supormos que Pat atribui a Mike uma “coisa” em relação aos ruivos. Essa “coisa” não é uma crença, não é uma atitude proposicional, mas se revela nas suas atitudes: um comportamento mais agressivo e hostil em relação aos ruivos, uma dificuldade maior para elogiar um ruivo e por aí vai... Pat pode estar mais correta em relação a Mike do que ele próprio sabe, pois a atitude dele não pode ser pinçada por meio da atribuição de uma crença isolada, nem por ele mesmo nem pelo intérprete. Seu comportamento está enraizado em um conjunto de crenças das quais ele não tem muita consciência e que não se sujeitam a codificação linguística. Os preconceitos influenciam o comportamento das pessoas de uma forma muito complexa, frequentemente sem que elas próprias possam perceber. Justamente por isso, não podem ser capturados em forma de crenças isoladas.

Assim, Dennett rejeita a ideia de que as atitudes proposicionais possam ser selecionadas e atribuídas de modo discreto. Elas não podem, para ele, ser individualizadas. Além do mais, a atribuição de estados mentais não inclui apenas atitudes proposicionais, mas também estados não proposicionais. Outro ponto: ao atribuir estados mentais a alguém, o intérprete pode fazê-lo de modo proposicional, por exemplo, “creio que João creia que choverá em breve”, como também pode fazer de modo não proposicional. Isso vale tanto para a atribuição de estados mentais proposicionais quanto para a atribuição de estados não proposicionais. É uma consequência de não se considerar a psicologia de senso comum como uma teoria.

Uma consequência dessa rejeição é que não se podem localizar estados cerebrais correspondentes a crenças isoladas, assim também não se encontram estados funcionais de um sistema correspondentes a atitudes proposicionais específicas. Dennett (2013b) apresenta uma situação para mostrar isso. Imagine que um cientista “neurocriptógrafo” implante na mente de um sujeito a seguinte crença falsa: “tenho um irmão mais velho que vive em Cleveland”. Então, esse sujeito diz isso aos seus amigos em um bar, mas logo é questionado: - “Como ele

se chama?” Rapidamente, o sujeito irá perceber que sua crença era falsa, pois não é possível inserir uma crença solta que não combine com as demais crenças daquele sujeito. Uma crença falsa depende, portanto, de uma complexidade de ajustes no conjunto de estados mentais daquele sujeito.

3.3. O método heterofenômico

Para o estudo das mentes humanas, Dennett propõe o desenvolvimento de um método que tem como material os relatos dos sujeitos acerca de seus estados internos. Para que isso funcione, o sujeito deve ser levado a sério, deve ser considerado como um sujeito racional, ou seja, o experimentador deve adotar a postura intencional. Se o experimentador partir da dúvida acerca de se o sujeito é ou não um zumbi, não será possível a aplicação do método. No mais, Dennett recomenda que sejam adotadas o mínimo de pressuposições possível (nesse ponto, o método se parece com o da tradução radical, quando o tradutor parte de um estado inicial minimamente informado acerca dos significados do discurso a ser traduzido).

O discurso do sujeito acerca de suas experiências internas é uma forma de descrevê-las. É a única descrição possível do sujeito acerca de si mesmo. Porém, é possível fazer descrições de um sujeito a partir de perspectivas externas e que também podem estar corretas. Em outras palavras, a descrição do sujeito e as descrições de seus interlocutores acerca dele são complementares na aplicação do método. Existem inclusive, formas de descrever o que se passa na vida subjetiva de alguém que não estão disponíveis para o próprio sujeito: por exemplo, a descrição do que acontece no seu cérebro a nível neuronal quando ele entra em um ambiente completamente branco¹⁰². Isso não significa que Dennett (2005b) desconsidere a autoridade de primeira pessoa acerca de seus estados mentais. Ao contrário, a aplicação do método heterofenômico consiste em reconhecer a autoridade do sujeito para falar sobre seu ponto de vista acerca do que acontece consigo. “Nós acreditamos em você para dizer como lhe parece; nós o constituímos como uma autoridade acerca disso” (DENNETT, 1982, p. 177, tradução nossa). Há coisas sobre o sujeito que ele mesmo não pode saber e precisam ser estudadas por outras vias, mas há coisas que só o sujeito pode saber. Só que essas coisas, na concepção de Dennett, poderão ser compartilhadas se esse for o desejo daquele sujeito.

¹⁰² Vale notar que há experiências das quais não podemos falar, mas elas nos provocam outras reações físicas ou comportamentais. Ao pensar isso, olhei pela janela e reparei que o céu está colorido de rosa, amarelo e azul. É lindo, está diante dos meus olhos, mas eu não teria podido falar sobre isso a menos que tivesse parado um pouco para prestar atenção ao que está acontecendo. Ao mesmo tempo, enquanto observo o céu, outras experiências passam por mim e posso mesmo dizer que meu corpo reage a elas, mas não posso falar sobre elas, pois não as percebo.

Você pode perguntar a si mesmo o que é a sua experiência subjetiva e ver o que você quer dizer. Então você pode decidir endossar sua própria declaração, acreditar nela, e então perseguir as implicações desse credo. Você pode fazer isso falando consigo mesmo em voz alta, falando silenciosamente consigo mesmo ou “apenas pensando” consigo mesmo sobre o que você está experienciando atualmente.

Essa é a extensão do seu acesso à sua própria experiência, e ela não difere muito do acesso que outra pessoa pode ter a essas experiências - *suas* experiências - se você decidir torná-las públicas com seu relato. (DENNETT, 2017b, p. 350-351, tradução nossa).

Essa compreensão de Dennett tem a ver com o fato de que, para ele, a mente humana é um tipo de mente muito específico, por ser linguística e cultural. A descrição de primeira pessoa acerca da sua fenomenologia é um material crucial para se compreender a subjetividade. Ocorre que para Dennett a descrição feita pelo sujeito acerca de seu mundo íntimo é uma forma de descrição, sendo que outras abordagens descritivas seriam possíveis. Tais abordagens, reunidas, poderiam trazer uma compreensão mais ampla e completa do que seja a mente.

Dennett denomina esse método de “heterofenomenologia”. Há a possibilidade de uma pessoa se enganar acerca das suas próprias experiências subjetivas, isso não precisa ser descartado¹⁰³, mas é uma possibilidade muito pequena. Além disso, é necessário que o intérprete atribua ao sujeito crenças, desejos e demais estados mentais para que os sons emitidos façam algum sentido e sejam compreendidos como algo que ele *queria* dizer. Para ele, ao interpretarmos os relatos do sujeito, permitimos que eles constituam o universo a ser pesquisado a partir da interpretação que fazemos daqueles relatos. É um universo ao mesmo tempo subjetivo e intersubjetivo, pois se constrói por meio do discurso. Então, embora o sujeito seja uma autoridade acerca de seus estados mentais, ele não é a única autoridade.

A consciência, nessa perspectiva, não é algo inacessível a uma perspectiva científica. O método heterofenomenológico é proposto como uma alternativa para essa abordagem. O método é detalhado inicialmente em 1982, mas permanece sendo apresentado como uma proposta para o estudo da consciência até no livro mais recente (2017b). Primeiramente, ele propõe que um grupo de sujeitos seja selecionado: “nós, os seres humanos mais ou menos normais e inteligentes” (DENNETT, 1982, p. 159, tradução nossa), pois é onde costumamos pensar que a consciência esteja. Em seguida, ele propõe que todo o estudo seja restrito àquilo que seja possível observar, a partir de fora, acerca do sujeito e que seja possível descrever objetivamente, com tantos detalhes quanto possível.

¹⁰³ Um exemplo bastante significativo é a prosopagnosia - uma desordem na qual o sujeito deixa de reconhecer os rostos das pessoas, inclusive de pessoas familiares, embora tenha a visão normal e a capacidade de reconhecer os objetos também.

O primeiro foco, segundo Dennett, será no texto composto das palavras da linguagem do sujeito. Essa etapa da aplicação do método consiste em uma passagem das propriedades físicas e acústicas dos sons emitidos para uma transcrição das palavras daquele sujeito. Nessa passagem, algumas interpretações precisarão ser feitas diante de sons inaudíveis ou palavras cortadas. Para que isso seja feito, é preciso uma série de pressuposições acerca da linguagem que está sendo falada e das intenções daquele sujeito. Aqui, o método heterofenômico mostra novamente suas semelhanças com o método de interpretação radical de Davidson (embora Dennett não o mencione nesse artigo, menciona a tradução radical de Quine). A etapa seguinte passa a exigir ainda mais pressuposições. Após a transcrição, é preciso haver a interpretação, mas esta etapa só é possível, segundo Dennett, se consideramos que aquele texto possui uma interpretação intencional. É preciso, então, atribuir crenças e desejos àquele sujeito para que seja possível interpretar o que ele diz. Porém, muito daquele discurso também é relativo ao que o sujeito acha que seu interlocutor (o experimentador que está aplicando o método heterofenômico) espera dele. São atribuições de crenças sobre atribuições de crenças, o que inevitavelmente poderá levar diferentes experimentadores a interpretações rivais do mesmo discurso.

Uma vez que Dennett havia proposto inicialmente que poucas pressuposições deveriam ser feitas e, em seguida, uma série de pressuposições se tornam necessárias, ele precisa deflacionar o método. Para isso, recorre a uma analogia com o método de interpretação de um texto de ficção. O fato de o texto ser de ficção não é um empecilho à sua interpretação. Porém, para interpretá-lo, é necessário fazer uma série de pressuposições além daquilo que está escrito no próprio texto, mas que podem ser consideradas aspectos objetivos daquele universo ficcional. “Todos os intérpretes concordam que Holmes era mais esperto que Watson; na completa obviedade repousa a objetividade” (DENNETT, 1982, p. 165, tradução nossa). O autor enfatiza que um texto que não seja de ficção pode também ser interpretado da mesma forma. Pode-se interpretar a biografia da rainha da Inglaterra sem que seja necessário ao intérprete responder à pergunta sobre se a rainha da Inglaterra realmente existe.

Há outra analogia usada por Dennett para esclarecer o método heterofenômico: a analogia com o trabalho de um antropólogo que pretende estudar a religião de uma tribo indígena. Nesse caso, o antropólogo não precisa se converter àquela religião, mas poderá se manter agnóstico acerca das referências dos discursos dos nativos. O antropólogo não precisa saber se aqueles relatos acerca da religiosidade daquelas pessoas possuem referências, não precisa acreditar naquela religião. Seu trabalho independe disso. Outro ponto importante é que o antropólogo precisará tomar os membros da tribo e, em especial, os seus sacerdotes, como

as maiores autoridades acerca do assunto. Também precisará atribuir aos discursos daquelas pessoas sinceridade no sentido em que elas provavelmente estão falando sobre algo em que acreditam.

A comparação com o trabalho do antropólogo revela duas características importantes do método proposto do Dennett: primeiramente, que o experimentador deve buscar, na medida de suas possibilidades, manter-se *agnóstico* acerca do que o sujeito está dizendo sobre seus estados internos. Afinal, o que o sujeito está relatando é como lhe *parece* o seu mundo interno. A segunda característica é a da neutralidade na aplicação do método, no seguinte sentido: levar a sério o relato do sujeito não significa atribuir verdade àquilo que o sujeito diz. Em outras palavras, a possibilidade de o sujeito estar enganado em algumas de suas crenças não está cerrada. Outro aspecto relativo à neutralidade é que o método pode ser aplicado independentemente de se o experimentador aceita ou não a realidade das experiências subjetivas. Porém, para Dennett, se essa realidade não pode ser alcançada a partir desse método, é preciso manter-se agnóstico em relação a ela.

Nesse ponto, Dennett compreende o método como minimalista em termos metafísicos, por duas razões: a primeira é essa apresentada acima, isto é, que na aplicação do método o pesquisador não precisa atribuir uma existência concreta ao mundo heterofenomenicamente constituído de alguém. A segunda razão é que o material para se estudar a subjetividade se constitui a partir dos relatos do sujeito acerca de seus estados internos. Mas o cientista heterofenomenologista não esperará encontrar dentro do sujeito nenhum tipo de referência para aquelas frases proferidas. “Enquanto essa visão trata de um fenômeno que é dependente da capacidade de produção de texto de alguns organismos, ela não pressupõe que esse fenômeno real seja ele mesmo linguístico, feito de palavras ou frases na cabeça, por exemplo” (DENNETT, 1982, p. 176, tradução nossa). Como resultado disso temos uma espécie de indeterminação. Não há apenas uma maneira pela qual o sujeito possa expressar o que se passa dentro dele. A postura a ser adotada pelo heterofenomenologista será de compreender o discurso como uma representação parcial do mundo fenomênico.

Na concepção de Dennett (1991b) é muito pouco provável que haja referência concreta no cérebro para aquilo de que o sujeito fala. A mente cultural é feita de abstrações e nela se encontram os sonhos, as alucinações e as fantasias, não há objetos internos aos quais a mente se direcione e sobre os quais construa os seus discursos. O método heterofenômico estuda o sujeito de um ponto de vista estritamente semântico. Por outro lado, é preciso haver uma base material adequadamente estruturada, com uma arquitetura apropriada e materiais eficazes para que uma mente possa existir. Essa base também precisa ser compreendida para

que se possa esclarecer o mental. A constituição do mundo heterofenômênico a partir dos relatos de um sujeito é uma parte importante da história, mas não é toda a história.

3.3.1. O mundo heterofenômênico de um robô

Todos os recursos disponíveis para as ciências cognitivas devem ser usados para que se possa compreender o mental - pensa Dennett - e o método heterofenômênico seria um desses recursos. Para o autor, nós fomos projetados pelo processo de seleção natural (arquitetura sem arquiteto inteligente) para ter certa percepção do que se passa dentro de nós, porém, outras histórias também poderiam ser contadas, se tivéssemos sido projetados de outra maneira. A explicação que temos para nossos estados internos é tão contingente quanto o próprio processo evolutivo. Um mesmo comportamento pode ter explicações divergentes, mas com idênticas propriedades explicativas, sob diferentes perspectivas. Por isso, embora haja uma perspectiva sobre a qual apenas o sujeito poderia falar, há também os pontos de vistas dos outros sujeitos e das diversas disciplinas científicas.

Para exemplificar isso, Dennett constrói um experimento de pensamento em que recorre ao robô Shakey. No início dos anos 1970, pesquisadores do Instituto de Pesquisa de Stanford desenvolveram esse que foi considerado a primeira inteligência artificial. Ele era equipado com uma câmera, sensores de colisão, um detector de distância e era conectado a um computador externo a ele, com um monitor, por onde os programadores trabalhavam. Shakey habitava ambientes fechados contendo objetos como caixas, pirâmides e calços. Seu trabalho era manipular esses objetos conforme as diretrizes recebidas. O programador comandava o robô à distância, por meio de um sistema de rádio e utilizava o monitor para observar os movimentos e para ter acesso ao processamento das informações por parte do robô.

Dennett (1991b) pede para supormos que Shakey seja equipado de tal forma que possa explicar como consegue identificar as formas para executar sua tarefa de manipulação dos objetos. Nesse caso, haveria diversos níveis de explicação possíveis a Shakey. Por exemplo, Shakey poderia dar uma explicação longa e detalhada usando linguagem de programação para explicar cada um dos subprocessos ocorridos internamente. Poderia também dizer que fez uma identificação dos contornos e dos vértices da figura, o que teria lhe permitido reconhecê-la. Poderia dizer que simplesmente teve a percepção de que algumas coisas são pirâmides e outras são caixas, mas não sabe dizer como foi que chegou a essa conclusão (essa resposta provavelmente se pareceria com as respostas de um humano diante da mesma tarefa).

Por último, segundo Dennett, o programador de Shakey poderia ter tido qualquer ideia mirabolante para que Shakey pudesse dar essa explicação. Por exemplo, o robô poderia ter sido feito para responder que fez o seu homúnculo mental se sentar sobre as figuras. Se o homúnculo escorregasse, um calço era identificado, se ele não conseguisse sentar, seria uma pirâmide, mas se o homúnculo conseguisse se sentar e ficar, significa que seria uma caixa.

Seguindo com o experimento de pensamento, Dennett pede para supormos que alguém com conhecimentos em programação verifique junto ao monitor todos os processos e subprocessos ocorridos em Shakey permitindo-lhe a identificação das figuras. Então, essa pessoa começa a dar uma explicação daquilo que ocorre internamente ao robô, mas a que o próprio robô não tem acesso. De acordo com Dennett, o robô talvez se rebelasse, dizendo que os estados internos *reais* dele são os que só ele pode ter acesso de um ponto de vista subjetivo. Nesse caso, Shakey estaria agindo da maneira como muitas pessoas tenderiam a agir diante da mesma situação. Mas isso também é um material na aplicação do método.

No caso de Shakey, por ele ser um robô, parece claro que a explicação do programador é tão ou mais importante do que a do próprio Shakey. Isso fica nítido para nós porque sabemos quem foi que o arquitetou e para quais propósitos. Nós não fomos arquitetados por uma inteligência que tivesse propósitos tão específicos ao nos fazer falar sobre nossos mundos internos. Mesmo diante dessa grande diferença entre nós e Shakey, podemos entender esse experimento de pensamento como uma analogia para compreendermos a proposta de Dennett de que, embora os relatos do sujeito acerca de seu ambiente interno sejam uma fonte importante de pesquisa, as demais fontes de pesquisa também são muito importantes e, eventualmente, poderão esclarecer aspectos do mental que os relatos do sujeito não são capazes de esclarecer.

Outro ponto de analogia é que, conforme Dennett (1991b, p. 96-97, tradução nossa), “Você não é autorizado sobre o que está acontecendo com você, mas apenas sobre o que parece estar acontecendo em você”. Isso que parece, para o sujeito, é o “como é ser” [“what is it like to be”] (NAGEL, 1974/2005), mas há outras coisas ocorrendo naquele sujeito sobre as quais o sujeito não é a maior autoridade. Nesse sentido, o ponto de vista do sujeito acerca do que está acontecendo consigo se torna uma abordagem superficial dos processos mentais. Assim como Shakey não conhece os detalhes de sua programação, nós não conhecemos os processos e subprocessos ocorridos em nossos próprios sistemas cognitivos.

“Essa metodologia de terceira pessoa, apelidada heterofenomenologia (fenomenologia *de outro*, e não de si mesmo), é, eu aleguei, o caminho sólido para levar o ponto de vista de *primeira* pessoa tão seriamente quanto ele pode ser levado” (DENNETT, 2003c, p. 1,

tradução nossa), Por isso, os diversos níveis de explicação possíveis devem ser levados em consideração, não apenas a explicação que o sujeito possui dos seus estados mentais, como também as explicações advindas de pesquisas realizadas nas mais diversas áreas científicas para o estudo do mental. Porém, conforme o autor, dentre as diversas explicações possíveis, somente aquela na qual o sujeito fala sobre seus estados internos é a que leva a subjetividade a sério sem deixar de adotar o ponto de vista de terceira pessoa.

A heterofenomenologia não parte das experiências conscientes para explicá-las, mas parte daquilo que está disponível como dado para quem aplica o método, dados objetivos, portanto. O heterofenomenologista precisa começar seus estudos pelos relatos verbalmente expressos do sujeito. Porém, é possível que o sujeito diga que suas crenças são sobre coisas subjetivas inefáveis. Segundo Dennett (2003c), nesse caso o heterofenomenologista não deve mudar seu foco de pesquisa para as coisas ditas inefáveis. Seria mais proveitoso para ele tentar entender o motivo pelo qual o sujeito alega possuir crenças sobre coisas inefáveis. As coisas inefáveis, segundo Dennett, não devem ser pressupostas para que se faça a pesquisa acerca da mente humana. O heterofenomenologista deve seguir seus estudos com os dados que possui.

Certamente, essa proposta seria rechaçada por quem parte da intuição de que a experiência subjetiva deve ser o ponto de partida no estudo do mental, ou que ela é o único material disponível para que o mental possa ser estudado. Dennett inverte esse processo. No estudo da consciência, o experimentador e o sujeito não podem ser a mesma pessoa o tempo todo, é preciso que outros sujeitos sejam estudados. A heterofenomenologia se opõe à autofenomenologia. Por isso, o experimentador não pode usar como dado suas próprias convicções acerca de suas experiências inefáveis. Também por isso, precisa estender a sua pesquisa a outros sujeitos. O material é o discurso do sujeito acerca das suas experiências. Com isso, está aberta para o pesquisador da heterofenomenologia a possibilidade de se explicar a subjetividade a partir de seus limites externos. Uma vez explicados esses limites, o mistério deixa de existir e ela não mais permanece presa dentro do sujeito.

3.4. Usos diferentes e complementares para o termo “consciência”

A compreensão de consciência de Dennett não é sempre a mesma. Quando o autor está tratando do modelo dos rascunhos múltiplos (1991b) ou do modelo da fama no cérebro (2005b), tem uma concepção muito específica e pouco convencional do termo. Nesses casos, os conteúdos conscientes não são estados subjetivos de acesso direto, mas são simplesmente

conteúdos que se tornam disponíveis para algum tipo de manifestação comportamental, inclusive conteúdos que não ficaram retidos na memória daquele sujeito e sobre os quais o sujeito não é capaz de falar.

[...] uma vez que uma “observação” particular de alguma característica tenha sido feita por uma porção do cérebro especializada e localizada, o conteúdo da informação então fixado não tem que ser enviado a nenhum outro lugar para ser re-discriminado por algum discriminador “mestre”. Em outras palavras, a discriminação não leva a uma representação da característica já discriminada para o benefício da audiência no Teatro Cartesiano - pois não há Teatro Cartesiano.

Essas fixações de conteúdo distribuídas espacial e temporalmente no cérebro são precisamente localizáveis tanto no espaço quanto no tempo, mas seus começos não marcam o começo da consciência de seus conteúdos. É sempre uma questão em aberto se qualquer conteúdo particular assim discriminado irá eventualmente aparecer como elemento na experiência consciente, e é uma confusão, como nós veremos, perguntar quando ele se torna consciente. [...]

Examinar esse fluxo [de consciência] em diferentes lugares e momentos produz diferentes efeitos, precipita diferentes narrativas a partir do sujeito (DENNETT, 1991b, p. 113, tradução nossa).

Essa primeira noção de consciência é anti-intuitiva porque, para Dennett, nosso sistema cognitivo não é capaz de lidar com a quantidade imensa de informações que nos bombardeiam em um tempo e um espaço muito pequenos. Nosso comportamento é, para ele, guiado por eventos aos quais nem sempre temos um acesso subjetivo privilegiado. Porém, como ele pensa que as informações se submetem a constantes processos de revisão no cérebro, uma história pode ser contada no momento em que o sujeito parar para se observar. É então que uma história é produzida a partir de uma seleção e edição das diferentes informações disponíveis para a versão final. Como ocorrem inúmeros eventos simultaneamente e nossos cérebros processam quantidades imensas de informação, o que se torna disponível para que uma história seja contada em um momento não está disponível em outro momento. Conseqüentemente, se alguém for contar a história de seu dia anterior, essa história será diferente da que seria contada pela mesma pessoa acerca das mesmas situações em um outro momento.

Outra consequência a ser tirada daí é que, na concepção de Dennett, apenas uma pequena parte da informação processada no cérebro poderá alcançar a consciência. Daí se pode tratar da sua segunda concepção do termo “consciência”, ligada à primeira por ser uma consequência dela. Aqui, a consciência é compreendida como uma história que o sujeito narra para si mesmo ou outrem quando olha para dentro. Por isso, ela se torna característica dos seres linguísticos. Então, embora a intencionalidade, a inteligência e até a racionalidade sejam

altamente disseminadas (tanto na natureza quanto na cultura), apenas os seres linguísticos seriam conscientes.

Vale a pena notar que essa compreensão da consciência como a história que é narrada no momento e local em que o sujeito dá uma breve olhadela para dentro vem de modo harmônico com as raízes interpretivistas do autor, quase como uma consequência natural da junção de suas bases filosóficas com os desenvolvimentos em psicologia, biologia e inteligência artificial que ele busca incorporar. A história que é narrada poderia ser outra em outro momento, não há a narrativa definitiva. Essa narrativa é editada, pois precisa se compatibilizar com uma série de outros estados mentais, crenças e desejos. Mas à versão dennettiana do interpretivismo se acrescentam outros estados cognitivos, além das atitudes proposicionais, embora esses outros estados não entrem na versão final da história. A versão final do processo editorial é um material linguístico acerca de estados supostamente subjetivos.

No fim das contas, estamos mais uma vez diante de um tipo de indeterminação. O que é indeterminado aqui é a interpretação que o próprio sujeito faz de sua experiência interna. O holismo do mental está pressuposto. Porém, uma série de experimentos em psicologia cognitiva levam o autor a incluir em suas pesquisas não apenas os estados mentais linguísticos, mas também uma série de informações processadas no cérebro sem que o sujeito possa perceber. Os modelos dos rascunhos múltiplos e da fama no cérebro acabam sendo formas de desenrolar o fundo interpretivista com as pesquisas mais recentes do autor.

A consciência, nessa concepção, é algo que surge recentemente na história evolutiva. Depois que alguns animais passaram a utilizar a linguagem para se comunicar e depois que essa linguagem se tornou complexa, esses animais começaram a olhar para dentro usando os mesmos hábitos cognitivos utilizados para lidar com o mundo externo. Com isso, puderam construir histórias sobre seus estados internos, histórias essas que, na concepção de Dennett, constituem a consciência, chamada por ele de centro de gravidade narrativa. A consciência é uma abstração onde orbitam as diversas narrativas constituintes de um sujeito.

3.4.1. Ilusionismo

Para Dennett (2016), a consciência deve ser compreendida como uma ilusão. Ele considera a posição chamada de ilusionismo suficientemente deflacionada para ser uma contribuição filosófica às ciências cognitivas. A palavra “ilusionista” é geralmente usada para designar alguém que faz um truque e causa na plateia a impressão de estar diante de uma

espécie de quebra na ordem natural das coisas. Algo que parece mágico ou sobrenatural, contudo, é fruto de eventos naturais. Nada sobrenatural ou mágico está acontecendo, é apenas uma impressão. “Nós ilusionistas filosóficos também somos ilusionistas no sentido cotidiano - ou deveríamos ser, isso é, nosso *fardo* é descobrir e explicar *como a ‘mágica’ é feita.*” (DENNETT, 2016, aspas internas do original, tradução nossa). Dessa forma, a filosofia seria capaz de trazer uma explicação para o mental em conformidade com as pesquisas científicas, sem interferir nelas.

A mágica da consciência, segundo essa concepção, é resultado de uma série de mecanismos cognitivos subjacentes. Explicando-os, a mágica desaparece, mas a aparência de mágica permanece (FRANKISH, 2016). Por isso, segundo Dennett, o ilusionista filosófico não encontrará sua realização tão cedo, pois será preciso muito desenvolvimento científico até que esses mecanismos se esclareçam. Mesmo assim, ao apresentar a consciência como uma mágica, o ilusionista neófito contribui muito, pois elimina das pesquisas em cognição muitos caminhos infrutíferos.

O ponto importante, para o ilusionismo, é que embora a consciência possa parecer algo que não se submete às leis naturais, ela não é. Descobrir o truque é simplesmente descobrir uma sequência inusitada ou surpreendente de eventos físicos: “Nós ilusionistas filosóficos dizemos que antes de ser impulsivo tomando as teorias da consciência como um tipo ou outro de ‘mágica real’, você deveria explicar tudo isso como uma ilusão engendrada pela natureza” (DENNETT, 2016, p. 67, aspas internas do original, tradução nossa).

Pode-se observar que mesmo se as pesquisas científicas demonstrassem que a consciência não é produto da cultura, o pressuposto do ilusionismo não precisa ser abandonado. Nesse caso, os cientistas precisariam se voltar para os níveis mais básicos de explicação. “É uma possibilidade remota que nós tenhamos que recorrer à física quântica ou a múltiplos universos para dar conta de um pouco de mágica incompreensível, mas primeiro, vamos tentar a rota conservadora” (DENNETT, 2016, p. 4, tradução nossa). Em outras palavras, Dennett compreende a consciência como fenômeno linguístico, portanto cultural, mas o ilusionismo independe dessa tese, por ser uma tese mais fraca.

3.4.2. *A mente cultural dos memes*¹⁰⁴

A noção de meme usada atualmente no sentido comum remete atualmente a uma imagem, um pequeno vídeo ou animação ou qualquer coisa que tenha um caráter imagético e muito repetido na internet, frequentemente com tema humorístico. Porém, o termo foi cunhado por Dawkins em 1976. Nessa época, embora a internet já fosse usada nos meios acadêmicos e militares em alguns países, Dawkins cunhou o termo com outro sentido, do qual o sentido comum se deriva.

Em evolução orgânica, Dawkins defende o ponto de vista do gene. Segundo ele, o gene é a unidade de seleção que faz o processo evolutivo seguir adiante, mas com o único propósito de produzir bons veículos para sua replicação. Por isso, desde que a reprodução genética possua boas chances, com fecundidade, fidelidade e longevidade, não importa as consequências para o portador. Nesse sentido, fenótipos que parecem mal adaptativos para um organismo individual podem ser bons veículos para a disseminação dos genes.

O gene é apresentado como uma unidade replicadora. Mas Dawkins amplia essa noção de unidade replicadora. Ao final de seu livro, constrói a hipótese especulativa de que a *cultura* também esteja à mercê de supostas unidades replicadoras, os memes. O intuito delas é se replicar, independentemente de serem boas ou más para a sobrevivência de seus veículos. Na época, os principais veículos de memes eram os seres humanos, mas havia também os livros, artefatos culturais diversos, músicas, revistas etc. Atualmente, a internet é um veículo que aumentou enormemente a capacidade de reprodução dos memes. Mas aqui, não estamos falando apenas dos memes da internet, sim de tudo que se reproduz culturalmente.

Segundo a hipótese da memética, a cultura é um produto dos memes. Eles se reproduzem em ambientes culturais, que vão desde as mentes humanas até o ambiente da internet. Assim como se reproduzem na cultura, também são capazes de modificar seu ambiente, tornando-o progressivamente mais favorável à sua própria replicação. A internet seria um exemplo de item cultural criado pelos memes para aumentar sua fecundidade, fidelidade e longevidade.

A memética traz consigo um descolamento entre os processos de evolução cultural e biológica, com interessantes consequências. Em primeiro lugar, torna-se possível relacionar

¹⁰⁴ Em Fagundes (2009), pode-se encontrar uma explicação da evolução memética como uma aplicação do algoritmo evolutivo, conforme exposto por Dennett. Porém, era um momento em que eu estava mais convencida em relação à memética. Também era um momento em que a própria memética estava mais abstrata e especulativa. Apesar disso, Richerson e Boyd (2005) já haviam escrito um livro sobre coevolução gene/cultura que tem muito em comum com a memética, mas traz os estudos para um plano mais concreto, envolvendo pesquisas populacionais acerca da cultura.

cultura e evolução sem a necessidade de buscar uma explicação biológica para aspectos culturais¹⁰⁵. Em segundo lugar, torna-se possível compreender porque muitas vezes somos culturalmente guiados a nos comportar de modos que parecem prejudicar nosso bem estar corporal, nossos relacionamentos e mesmo nossa sobrevivência e reprodução. É que, além de termos que nos preocupar com nossos “eus” biológicos, precisamos também cuidar dos “eus” culturais.

A consciência humana é distinta de todas as outras variedades de consciência animal por ser um produto em larga medida da evolução cultural, a qual instala um subsídio de palavras e muitas outras ferramentas de pensamento em nossos cérebros, criando assim uma arquitetura cognitiva diferente das mentes de baixo para cima dos animais. Abastecendo nossas mentes com sistemas de representações, essa arquitetura equipa cada um de nós com uma perspectiva - uma ilusão do usuário¹⁰⁶ - a partir da qual nós temos um acesso limitado, enviesado, aos trabalhos dos nossos cérebros [...]. (DENNETT, 2017b, p. 370, tradução nossa)

A consciência humana, segundo o autor, envolve compreensão de cima para baixo. Como ela foi construída por um processo evolutivo, vem de baixo. Antes, havia competência cognitiva sem que houvesse compreensão. A compreensão surge depois e pode inverter o sentido da cognição. Essa é, para ele, uma inversão proposta por Darwin. Nós podemos compreender o processo evolutivo, mas não foi preciso haver uma compreensão desse processo para que ele ocorresse. A própria natureza, para Dennett, possui competência sem compreensão. “A compreensão - *nosso* tipo de compreensão - só se torna possível com a chegada bem recente de um novo tipo de replicador evolucionário - entidades informacionais culturalmente transmitidas: memes” (DENNETT, 2017b, p. 175, grifo do original, tradução nossa).

Para a aplicação do método heterofenômico¹⁰⁷, Dennett sugere que o pesquisador se mantenha agnóstico acerca de quais seres possuem ou não possuem consciência. Porém, ao compreender a consciência como uma ilusão das mentes linguísticas, Dennett acaba por se

¹⁰⁵ Isso evita o erro do darwinismo social que, nos anos de 1870, buscou aplicar os conceitos darwinistas à sociologia de forma desastrosa. O darwinismo social buscava justificar injustiças sociais apelando para conceitos darwinistas, levando à criação de políticas públicas que buscariam uma suposta evolução da espécie humana. Foi uma má compreensão dos conceitos darwinistas, mas foi usada, inclusive, para justificar o nazismo. É compreensível que hoje haja uma grande desconfiança nas ciências sociais em relação a qualquer abordagem darwinista da cultura. Porém, nada há nas características biológicas de um ser humano que vá determinar suas características culturais, pois os processos são separados.

¹⁰⁶ Ilusão do usuário é o que vemos diante da tela do computador. Não vemos como ele foi programado, nem como funciona. Ou seja, não conhecemos os mecanismos adjacentes, o que nos provoca a ilusão de estarmos diante de uma máquina incrivelmente inteligente. Mas não é uma mágica real. Assim como a consciência humana, é só ilusão.

¹⁰⁷ A heterofenomenologia é apenas uma proposta de método para o estudo da consciência e pode ser adotada independentemente de se abraçar ou não a memética.

comprometer com a ideia de que apenas para nossa espécie¹⁰⁸ faz sentido a pergunta “como é ser, para mim, o que sou?”. A consciência humana, para ele, seria uma ilusão linguística. Dado que a linguagem é um aspecto da cultura, Dennett compreende a consciência como um produto dos memes.

Há, para ele, um “princípio biológico fundamental de distinguir o eu [self] do mundo, lado de dentro do lado de fora” (1991b, p. 414, tradução nossa). Qualquer organismo necessita disso para se abrir aos nutrientes e se fechar às ameaças externas. Algumas espécies foram capazes de ampliar seus limites territoriais, como aranhas, que constroem teias. Essas ampliações dos limites de uma espécie foram chamadas de fenótipo estendido por Dawkins. Às vezes, o indivíduo não é capaz de realizar sozinho essa ampliação de limites, como no caso das formigas que trabalham em grupo. Porém, segundo Dennett, o caso humano é diferente porque o nosso próprio “eu” é um fenótipo entendido, nós o construímos por meio de narrativas contadas por nós sobre nós mesmos, mas também pelas narrativas acerca de nós contadas pelos outros. Isso ocorre, na concepção de Dennett, porque as outras espécies reproduzem genes, enquanto nós reproduzimos memes. Assim, nenhuma espécie possui um fenótipo tão extenso quanto nós.

Um “eu” [self], de acordo com minha teoria, não é qualquer ponto matemático antigo, mas uma abstração definida pelas miríades de atribuições e interpretações (incluindo autoatribuições e autointerpretações que compuseram a biografia do corpo vivo que é Centro de Gravidade de Narrativa (DENNETT, 1991b, p. 428, tradução nossa).

Para Dennett, nós estamos sujeitos a uma série de ilusões cognitivas que criam em nós a impressão de que estamos diante de um espetáculo no Teatro Cartesiano. Os *qualia*, para ele, são apenas estados disposicionais complexos do cérebro e o “eu” humano é uma criação linguística, não um expectador interno. Por essa razão, os estados supostamente subjetivos podem ser estudados a partir da adoção do método heterofenômico. A consciência humana, na concepção de Dennett, é totalmente diferente dos sistemas cognitivos das outras espécies. Esse salto provém do fato de sermos animais culturais, veículos de dois tipos de replicadores: genes e memes.

¹⁰⁸ Embora essa possibilidade permaneça remotamente aberta para as inteligências artificiais do futuro.

3.4.3. *O que seria um meme?*

Foi visto¹⁰⁹ que, segundo Dennett, para o algoritmo evolutivo entrar em cena, é necessário haver três elementos: variação, retenção e seleção. A questão aqui é: o que varia, é retido e é selecionado? Não é preciso haver uma unidade replicadora? Conforme Dennett (1990), os memes são “unidades memoráveis distintas”, como arco, roda, usar roupas, triângulo retângulo, movimentos artísticos... Porém, não fica claro quais são as unidades replicadoras. O problema, quando se trata de cultura, é que estamos lidando com complexos muito difusos, como poderíamos isolar uma unidade cultural e considerá-la como replicadora? No caso dos memes da internet, o meme pode ser simplesmente aquela figurinha que as pessoas replicam em suas redes sociais. Mas isso deixa de levar em consideração a grande quantidade e a complexidade das relações que esse meme replicador mantém com vários outros aspectos do entorno histórico e social sem os quais o meme da internet não poderia ser compreendido¹¹⁰.

Poderia-se argumentar que os organismos são extremamente complexos, especialmente se forem comparados aos genes e que ainda assim os organismos podem ser vistos (numa perspectiva do gene) como veículos de genes. Aqui, deve-se considerar que os cientistas genéticos são capazes de manipulá-los, o que sugere, mesmo para uma leiga, que há uma clareza em relação à unidade que está se reproduzindo, a tal ponto que é possível interferir artificialmente e alterar o curso das relações entre essas unidades. No mais, o gene é uma unidade estável, ao contrário do meme, pois os processos culturais mudam em uma velocidade extremamente acelerada.

Além disso, por mais complexo que seja um organismo, ele possui limites, são limites porosos, mas permitem uma separação entre os ambientes interno e externo do organismo. No caso da cultura, é muito difícil, se não impossível, encontrar esses limites. A memética tem enfrentado muitas críticas há bastante tempo. As críticas tendem a enfatizar as dissemelhanças entre os memes e os genes.

A maneira como Blackmore enfrenta essa crítica é dizendo que a analogia entre memes e genes é frouxa e que o próprio Darwin pôde construir sua teoria mesmo sem ter uma imagem da evolução baseada em genes. O aspecto importante da analogia entre genes e

¹⁰⁹ Ver: seção 1.4.1.

¹¹⁰ Uma cultura globalizada permite um fundo de conhecimentos compartilhados cada vez mais amplo e capaz de envolver mais pessoas. Isso poderia ser usado como argumento pelos defensores da memética: que a globalização é obra dos memes para benefício próprio. Mas esse argumento não é muito convincente, pois parece depender da aceitação prévia dos memes, é especulativo demais.

memes, segundo ela, é que ambos carregam informações. Segundo ela, “[n]ós podemos percorrer um caminho com os princípios gerais da seleção memética sem entender os mecanismos dos quais ela depende” (BLACKMORE, 1999, p. 56, tradução nossa).

Porém, a teoria de Darwin passa a ganhar grande aceitação científica após a síntese com o trabalho de Mendel, ocorrida apenas em 1942. Mendel começa a fazer suas pesquisas com cruzamentos de ervilhas em 1856 e Darwin publica sua teoria da seleção natural em 1858. Ou seja, quase um século se passou até que houvesse a síntese. A síntese compreende os genes como elementos de herança ao longo de gerações e traz um bojo ao trabalho de Darwin por mostrar como o processo evolutivo pode se dar a partir da replicação dos genes.

Dizer que Darwin pôde desenvolver a sua teoria sem ter conhecimento dos mecanismos envolvidos no processo para dizer que a memética também pode se desenvolver sem que se conheçam os mecanismos envolvidos no processo de evolução cultural parece carregar uma promessa: de que a perseverança no estudo memética permitirá compreender com detalhes os mecanismos envolvidos no processo. Ora, mas esses mecanismos envolvem a descoberta da unidade replicadora. A promessa parece otimista demais para justificar a memética como campo de pesquisa.

Mais tarde (2017b), Dennett identifica as palavras como exemplos paradigmáticos de memes. À primeira vista, isso ajudaria a uma aproximação em relação às unidades replicadoras. “Palavras, eu argumentarei, são o melhor exemplo de *memes*, itens culturalmente transmitidos que evoluem por replicação diferencial - isto é, por seleção natural” (DENNETT, 2017b, p. 176, tradução nossa). Dennett argumenta, aqui, que as palavras são discretas e possuem reprodução fiel. Ele faz referência a uma comparação feita por Dawkins (2004, *apud* DENNETT, 2017b) entre palavras e genes. De acordo com Dawkins, palavras são menores que genes, mas assim como genes, podem se combinar de diferentes maneiras em diferentes contextos. Para Dennett uma característica importante das palavras é que elas têm um tipo de digitalização: os fonemas ou os sinais gráficos que carregam seu conteúdo cultural, aproximando os memes dos genes no quesito da digitalização.

Os exemplares [“tokens”] de “gato” não são réplicas *físicas* de seus ancestrais, mas eles são - nós poderíamos dizer - réplicas virtuais, dependentes de um sistema finito de normas pelas quais os falantes inconscientemente corrigem suas percepções e suas enunciações, e isso - não a replicação física - é o que é requerido para transmissão de informação de alta fidelidade (DENNETT, 2017b, p. 227, tradução nossa).

Outro exemplo de Dennett é o da literatura. Segundo ele, duas impressões diferentes, mesmo que feitas em diferentes fontes, são cópias do mesmo romance, assim como diferentes traduções para um idioma ou para idiomas diferentes. O fato de a tradução envolver compreensão, segundo ele, não impede a confiabilidade da transmissão cultural. Aqui, Dennett passa da palavra para algo bem mais complexo, o romance. A transmissão do romance, segundo ele, envolve *compreensão*.

Aqui, a crítica à noção de meme como uma unidade replicadora pode se apresentar, contrariamente ao que Dennett pretendia com seu exemplo. Acontece que, nesse nível, já se torna muito difícil determinarmos o que está sendo reproduzido. A compreensão de um romance envolve uma série de outros aspectos culturais. Quanto mais distante um leitor estiver daquela cultura onde o romance foi construído, mais diferente será sua compreensão daquele romance em relação às pessoas que viviam no mesmo ambiente cultural do autor. Essas diferentes compreensões não impedem que pessoas das mais diversas culturas leiam o romance nas mais diversas línguas e aleguem que leram aquele mesmo romance. Porém, impedem que se determine qual a interpretação correta daquele romance, colocando-nos diante de um caso de indeterminação que, provavelmente, impossibilita a determinação da unidade replicadora.

Mas será que esse problema ainda permanece quando descemos para o nível da replicação das palavras? Sim, o problema permanece e talvez até se agrave. Uma palavra isolada depende de uma série de outras informações para que se possa ser compreendida. Seu significado não existe isoladamente de outros significados. Caso contrário, teríamos apenas ruídos ou rabiscos em um papel. O sinal de “gato” não fará qualquer sentido se eu não tiver me envolvido em relações de triangulação na infância que me ensinaram palavras por meio de ostensão. Uma palavra emitida não é uma palavra solta. Se eu não souber em que idioma estamos ambos falando, se eu não tiver como pressuposto que o emissor de “gato” possui crenças compartilhadas comigo, que ele quer se comunicar, e mais uma série de outras atribuições de atitudes proposicionais, “gato” se torna um ruído qualquer, sem nenhum sentido. Por isso, aparentemente, a palavra não é essa entidade replicadora discreta que pareceria à primeira vista.

Há duas lições provenientes das raízes interpretivistas de Dennett que são de difícil conciliação com a memética: o holismo dos termos mentais e a indeterminação da interpretação (e Dennett tem uma postura radical acerca da indeterminação). Isso ocorre porque, em uma posição como a interpretivista, nada na linguagem é concebido de modo isolado, mas sempre em relação com a mente e com o mundo. Se a perspectiva interpretivista

é levada adiante, como Dennett leva, a mente humana é sim um produto da cultura. Mas o problema da memética não é esse. Também não é conceber a cultura como um processo que segue sua evolução paralelamente e independentemente da evolução biológica, com pontos de intersecção entre elas¹¹¹. Aliás, esse é provavelmente um ponto forte da memética. O problema é a pressuposição de que haja um replicador, o meme, pois a cultura envolve complexos informacionais muito difusos, altamente conectados e emaranhados. Fica muito difícil encontrar o replicador para a memética. A memética parece carregar uma promessa muito difícil de se realizar e aparentemente incompatível com outros aspectos do pensamento de Dennett: a promessa de se descubra o aspecto discreto capaz de reprodução de alta fidelidade no meio da cultura.

3.5. Evolução, mente e linguagem em Dennett

O interpretivismo é algo bastante desenvolvido no pensamento de Dennett, especialmente na década de 1980, quando a influência de seu professor Quine se faz bastante presente e suas ideias estão próximas às de Davidson. As raízes dos interesses filosóficos de Dennett, contudo, começam a se estabelecer duas décadas antes, quando se torna importante para ele compreender o mental de modo compatível com as ciências empíricas e a partir de suas bases evolutivas.

Eu comecei a pensar seriamente sobre a evolução da mente humana quando eu era um estudante de pós-graduação em filosofia em Oxford em 1963 e não sabia quase nada sobre evolução nem sobre a mente humana. Naqueles dias, não se esperava dos filósofos que soubessem sobre ciência, e mesmo os mais ilustres filósofos da mente eram amplamente ignorantes dos trabalhos em psicologia, neuroanatomia, e neurofisiologia. (DENNETT, 2017b, p. XV, tradução nossa)

Dennett envereda pela trilha recém aberta do naturalismo. Para muitos, naquele momento, a filosofia precisa se aproximar da ciência para compreender a mente e qualquer outro aspecto da realidade a partir de seu papel na natureza. Desde então, o pensamento dele sempre esteve em diálogo com o que vem sendo produzido nas ciências cognitivas. Atualmente, segundo o autor, há centenas de jovens filósofos com formação interdisciplinar em ciências cognitivas, assim como há estudantes de ciências filosoficamente orientados.

Outro ponto do pensamento de Dennett que se estabelece desde seus primeiros livros é a ideia de que a mente se estuda a partir da adoção de uma determinada postura perante

¹¹¹ Aqui, surge uma dúvida para a qual não se procura aqui uma resposta: é possível pensar em um processo evolutivo aplicado a um contexto diferente do biológico, mas que não envolva um replicador discreto e estável?

alguém ou algo: a postura intencional. A mente, no fim das contas, se revela sob uma perspectiva de terceira pessoa, a perspectiva do intérprete. Não é possível tratar de algo que não possa ser estudado por uma perspectiva objetiva, segundo o autor. É nesse sentido que ele desenvolve o método heterofenômico. Sua teoria do mental também se enraíza no holismo das atitudes proposicionais e na indeterminação de Quine, levando-a de modo bastante radical.

A teoria dos sistemas intencionais de Dennett considera como intencional qualquer sistema cujo comportamento possa ser explicado ou previsto com sucesso por meio da atribuição de intencionalidade. Há várias ordens de intencionalidade. Nós somos sistemas intencionais de quarta ordem, no mínimo, pois precisamos disso para podermos nos comunicar pela linguagem articulada culturalmente aprendida dos humanos. Mas há sistemas intencionais de ordens inferiores. É importante frisar que Dennett recusa qualquer distinção entre intencionalidade original e derivada. Para Dennett, atribuímos intencionalidade aos outros da mesma forma como atribuímos a nós mesmos, e é assim que criamos a consciência tipicamente humana.

O termo “consciência” ganha algumas concepções diferentes ao longo do trabalho do autor, isso vai depender da distância a partir da qual a pesquisa é feita. Olhando para o sujeito muito de perto, chega-se a um ponto de vista inacessível ao próprio sujeito. Daí, os conteúdos conscientes são aqueles que se tornam disponíveis para manifestação comportamental, quer o sujeito tenha condições de falar sobre eles, quer não.

Afastando-se um pouco, contudo, Dennett pensa que o sujeito humano é capaz de construir uma narrativa acerca de seus estados internos. É uma espécie de ilusão que se cria quando o sujeito olha para dentro adotando os hábitos cognitivos que vinha usando para lidar com o mundo externo. Então, em outra concepção do termo, a consciência é constituída de narrativas do sujeito acerca de seus estados internos. Essas narrativas também são interpretadas pelo interlocutor daquele sujeito. O sujeito diz algumas coisas sobre si, mas só as pode dizer para alguém que precisará interpretar aquele discurso. No fim das contas, a interpretação do discurso do sujeito acerca de seus estados internos está sujeita à indeterminação.

A terceira concepção de consciência está bastante atrelada à segunda. Aqui, a consciência é compreendida como uma espécie de nicho ecológico criado pelos memes para sua reprodução. De acordo com Dennett, a cultura promove uma mudança na nossa arquitetura mental, criando a ilusão do usuário. Esse ambiente teria favorecido o próprio

processo de evolução cultural e a criação de “eus” humanos altamente expandidos e voláteis, mas com a aparência de integração.

As três concepções não são incompatíveis, isso depende da distância tomada em relação ao sujeito para que se possa estudá-lo, pois o nosso sistema cognitivo, para Dennett, funciona por meio de edições e o resultado final é uma espécie de integração de inúmeras informações fragmentadas recebidas por nós a todo instante, em uma velocidade muito maior do que aquela que somos capazes de monitorar. É por isso que criamos histórias aparentemente integradas para o que somos nós e o que são os outros sujeitos. Nós precisamos de uma distinção entre dentro e fora, ainda que ingênua, pois isso é importante para que possamos nos aproximar do que nos parece bom e nos afastar do que nos parece ruim. Isso foi importante para nossos ancestrais e, na concepção de Dennett, contribui para a formação da ilusão da consciência.

Aparentemente, contudo, há uma incompatibilidade entre a memética e alguns aspectos do interpretivismo. O interpretivismo é uma posição que olha para a mente e a linguagem, não para todos os aspectos da cultura. Porém, a mente e a linguagem estão sujeitas ao holismo no pensamento dennettiano. Sendo assim, ao menos os memes relativos a esses aspectos da cultura também se sujeitariam a esses dois fatores. Mas o que são os memes senão entidades replicadoras? E, quando se trata da mente, como isolar essas entidades das redes complexas formadas pelas atitudes proposicionais? O problema da memética é que ela fica muito volátil se não conseguimos isolar os replicadores da cultura. Mas se a mente é cultural e está sujeita ao holismo, isso nunca será possível. Nesse sentido, a memética provavelmente carrega promessas que não poderá cumprir.

4. COMO COMPREENDER OS ESTADOS MENTAIS

Para buscar uma compreensão do mental a partir de uma perspectiva interpretivista, é importante abandonar a ideia de que esse estudo deva ser feito a partir de um ponto de vista estritamente subjetivo. A ideia seria, então, buscar uma compreensão das relações entre o sujeito e o mundo físico e social. Para isso, a disciplina filosofia da mente não será compreendida como tendo nos *qualia* seu objeto central de estudos. As questões acerca de porque nossas vidas mentais parecem um universo inefável, intrínseco, de acesso direto e privado só do sujeito se torna um problema a ser confrontado a partir desse outro ponto de vista: das relações do sujeito com o mundo. É possível que, abordando os estados mentais de modo disposicional, essas questões acerca da consciência fenomênica acabem por se tornar mais tratáveis, essa é uma aposta de Dennett, conforme será abordado à frente. Por isso, a proposta é focar, ao menos em um primeiro momento, no estudo das atitudes proposicionais, mas tentando não ignorar os estados mentais não linguísticos.

Foram duas as classes de questões propostas para serem examinadas nesta tese: questões de tendência ontológica e questões de tendência epistemológica. Podemos, agora, retomar aquilo que foi proposto no início. O que se percebe, em uma abordagem que procura superar a dualidade entre objetividade e subjetividade, é que questões ontológicas e epistemológicas estão profundamente conectadas e não podem ser abordadas separadamente. Essa vinculação entre ontologia e epistemologia faz com que algumas das questões propostas inicialmente acabem por se diluir. Neste capítulo, procuraremos compreender melhor essa vinculação entre aspectos epistemológicos e ontológicos referentes aos estados mentais. Ao final, as questões propostas na introdução serão retomadas.

Qual é o status ontológico que se poderia atribuir às atitudes proposicionais? Há algo no mundo que tenha um status ontológico parecido? Se, numa abordagem interpretivista, as atitudes proposicionais são algo que se torna manifesto a partir da perspectiva de quem interpreta o comportamento alheio, então, faz sentido dizer que elas existem dentro do sujeito? A hipótese que se pretende defender é que os dois autores estudados, Davidson e Dennett, possuem respostas diferentes para a questão, sendo que Davidson defende um realismo um pouco mais forte que Dennett. Porém, as posições de ambos são brandas, nada de teses ontologicamente pesadas.

Espera-se que, no ponto em que estamos, os estados mentais não sejam pensados mais como algo que se encontra fechado dentro das mentes dos sujeitos. A questão não é, portanto, pensar sobre se há alguma referência *dentro* da mente do sujeito para nossas atribuições de estados mentais. As atitudes proposicionais e as atribuições de estados mentais formam um aspecto da realidade humana que não cabe em nenhum tipo de descrição física do cérebro. É preciso virar pelo avesso a concepção moderna de mente e passar a vê-la como algo que se constrói por intermédio de uma constante relação do sujeito com o mundo e os outros sujeitos. A subjetividade humana, em perspectivas como a de Dennett e a de Davidson, é construída a partir das relações de comunicação.

Aqui, Dennett e Davidson serão pensados como autores para os quais os aspectos ontológicos e epistemológicos se tornam conectados. Em Davidson, a linguagem promove essa conexão por meio da triangulação. Em Dennett, a necessidade que temos de atribuição de estados mentais é tão vital que, pelo fato de funcionar, confere realidade aos padrões de atribuições de estados mentais.

O capítulo está estruturado da seguinte maneira: na primeira seção, se pensará sobre o realismo de Davidson, um tipo de realismo fundamentado nas nossas relações de comunicação quotidianas. Em seguida, trataremos de como Dennett pensa a realidade das atribuições de estados mentais. Elas encontram sua força ontológica na indispensabilidade. A noção de indeterminação conduzirá a discussão entre Dennett e Davidson acerca do realismo, principalmente nas três primeiras seções. Na terceira seção, discutiremos o realismo moderado desses autores. A quarta seção tratará das relações entre os estados mentais e o mundo, pensando sobre a capacidade cognitiva dos organismos se direcionarem à objetividade. A última seção é dedicada às perguntas que haviam sido apresentadas na introdução do presente trabalho.

4.1. Davidson: a verdade das atribuições de estados mentais

Pensar sobre a ontologia das atribuições de atitudes proposicionais em Davidson deve pressupor o pensar sobre a verdade das crenças em geral, pois o mesmo argumento que dá suporte às nossas crenças sobre o mundo também dá suporte às nossas crenças sobre os estados mentais das outras pessoas. Trata-se aqui do argumento da triangulação.

Importante lembrar que, no contexto do pensamento davidsoniano, para podermos dar significado à linguagem, precisamos atribuir atitudes proposicionais aos outros. Não poderíamos sequer compreender nossas próprias crenças sem isso. Nesse sentido, em

Davidson, não há uma prioridade do conhecimento da própria mente em relação às outras mentes e ao mundo exterior (embora haja uma assimetria entre o conhecimento que temos de nossas próprias mentes e o conhecimento que temos das mentes alheias)¹¹². Além disso, “ter o conceito de crença é possuir uma certa capacidade, incluindo a capacidade de atribuir crenças às pessoas nas condições apropriadas - saber quando asserir coisas que, em português, podem ser traduzidas em sentenças da forma ‘a crê que p’” (STROUD, 2005, p. 105, aspas duplas do original substituídas por aspas simples). Crer e atribuir crenças são habilidades inseparáveis conforme o pensamento de Davidson, uma não existe sem a outra.

As nossas crenças, conforme o importante argumento davidsoniano, são majoritariamente verdadeiras porque são linguísticas e a linguagem se constitui da relação direta entre aprendiz, professor e mundo. O argumento da triangulação demonstra que, para aprendermos uma linguagem, precisamos atribuir crenças aos outros. Além disso, precisamos relacionar essas crenças atribuídas com a realidade compartilhada entre eles e nós. O fato de nos comunicarmos com sucesso, então, traz uma sustentação tanto às crenças sobre o mundo exterior quanto às crenças sobre o que se passa nas mentes dos outros sujeitos.

Mostrar a tendência à formação de um corpo de crenças majoritariamente verdadeiras a partir das relações quotidianas de comunicação é algo que tira da noção de “verdade” um caráter de universalidade sem lhe tirar a objetividade e, ao mesmo tempo, a torna alcançável a qualquer ser humano (independente das suas tendências filosóficas). A tendência à formação de crenças majoritariamente verdadeiras se constitui a partir de eventos locais, específicos, ocorridos nas vidas de todas as pessoas. Assim, a concepção de Davidson acerca da verdade se torna algo palpável. Com simplicidade e elegância, ele se desvia do ceticismo. Basta uma troca de olhares com alguém e um corpo difuso de crenças poderá ser formado, juntando-se esse corpo às demais crenças de alguém, a maior parte será verdadeira. Além disso, também haverá uma taxa de coincidência majoritária entre as crenças das pessoas. Deve-se manter em mente, diante disso, que a verdade de cada crença não é capturada de modo isolado, conforme a concepção holista defendida pelo autor - herdada por ele e Dennett - de Quine.

4.1.1. *Quais não são as concepções de verdade de Davidson*

São duas as concepções de verdade rejeitadas explicitamente por Davidson (1990/2002a, p. 79). A primeira é a concepção *epistêmica*, segundo a qual a verdade se restringe ao cognoscível, sendo portanto dependente das crenças. O autor condena esse tipo de

¹¹² Ver: seção 1.3.

concepção por ela abandonar a objetividade. Davidson não aceita a desconexão entre objetividade e subjetividade que acaba por prender a verdade na mente subjetiva, conduzindo ao relativismo. Ele defende que a verdade deve possuir alguma relação que a vincule à realidade circundante.

A outra concepção de verdade, chamada de realista ou *transcendental* também é rejeitada, pois a relação da verdade com a realidade não pode ser confirmada diretamente. “[N]unca seremos capazes de dizer quais de nossas crenças são verdadeiras” (DAVIDSON, 2000/2002c, p. 133). Nesse ponto, Davidson concorda com Quine (1960/2010) de que não é possível traçar uma linha demarcatória entre proposições analíticas e sintéticas. Assim, não podemos confrontar supostas proposições sintéticas isoladamente com os fatos. Isso torna a noção de verdade como correspondência algo vazio e inútil se for aplicada a proposições isoladas. Porém, as teorias da verdade como correspondência têm uma vantagem, na concepção de Davidson: a intuição de que há a verdade objetiva, ainda que, devido à indeterminação, não haja uma forma unitária de descrevê-la. A verdade objetiva é algo que independe das crenças das pessoas. Davidson não abre mão dessa noção, mas tenta alcançá-la por uma via diferente, uma vez que não existem crenças que possam ser isoladas para o confronto com os fatos.

O problema dessas duas concepções de verdade - a epistêmica e a transcendental - se pode perceber facilmente frente às ideias de Davidson: ambas provém de uma ruptura entre subjetividade e objetividade. A verdade epistêmica seria algo completamente dependente das crenças do sujeito, enquanto a realista seria independente ao ponto de ser inalcançável pelo sujeito. A verdade, para Davidson, é algo que nós podemos alcançar, mas ela não deixa de ser objetiva.

Conforme o pensamento davidsoniano, a noção de verdade deve surgir juntamente com a de objetividade no contexto da comunicação. Nosso entendimento de uma linguagem depende de sabermos as condições de verdade das frases. Mas isso depende de sabermos que a verdade é independente das nossas crenças, e que algumas de nossas crenças podem estar equivocadas. Assim, a compreensão da linguagem requer a noção de erro que, por sua vez, depende da verdade objetiva.

O fato de nos comunicarmos quotidianamente com as outras pessoas, nessa perspectiva, pressupõe a objetividade da verdade. É por isso que a noção de verdade objetiva promove uma via de mão dupla entre objetividade e subjetividade. A partir da compreensão da linguagem é que surge a exigência de que a verdade seja objetiva. A verdade objetiva, porém, não pode ser fígada por proposições isoladas. Sabemos uma série de coisas, temos

crenças majoritariamente verdadeiras, mas não temos como saber, acerca de crenças isoladas, se elas têm correspondência com o mundo.

Ao mesmo tempo em que a linguagem se torna um argumento para defender a objetividade e a independência da verdade em relação às nossas crenças, também imbui as próprias atribuições de atitudes proposicionais de objetividade. As atribuições de atitudes proposicionais podem ser verdadeiras ou falsas, e é preciso que nossas atribuições de atitudes proposicionais sejam majoritariamente verdadeiras para que a comunicação seja possível. Mas a comunicação é possível! As atitudes proposicionais alheias independem de nossas crenças, só se revelam a partir de um processo de interpretação. Esse processo permite que nós atribuamos legitimamente um conjunto de crenças majoritariamente verdadeiras aos outros e que estejamos também majoritariamente corretos em nossas atribuições.

Essa forma de compreender a verdade proposta por Davidson pode ser conciliada com a ideia da indeterminação. Aqui, Davidson reestrutura a noção de indeterminação da tradução como indeterminação da interpretação, abrandando-a em relação a Quine. De modo bastante superficial, pode-se apresentar a indeterminação da interpretação da seguinte forma: diante de duas interpretações rivais e igualmente preditivas do comportamento de um sujeito, não é possível escolher a melhor. Esse item da preditividade é o que abranda a concepção de indeterminação de Davidson, trazendo, ao final, certo grau de realismo às atitudes proposicionais de um sujeito.

4.1.2. *A analogia com os sistemas de medição*

A indeterminação, como compreendida por Davidson, está presente sempre que “um vocabulário seja rico o suficiente para descrever um fenômeno de mais do que uma maneira. Não importa se você diz que Sam está à esquerda de Susan, ou que Susan está à direita de Sam.” (1998/2005a, p. 316, tradução nossa). Estamos diante, portanto, da compreensão bastante cotidiana de que há diversas maneiras para se dizer alguma coisa. Isso possui consequências importantes para a posição do autor acerca da ontologia das atitudes proposicionais. Para ele, diferentes interpretações são, no fim das contas, apenas diferentes modos de dizer a mesma coisa.

Davidson (1997/2001d, p. 130 e segs.) propõe uma analogia entre as teorias da semântica formal e as da metrização fundamental. Ambas devem possuir um ou mais conceitos primitivos. No caso da metrização do peso, por exemplo, o conceito primitivo pode ser a relação de *ser ao menos tão pesado quanto*. A teoria mostra que, atribuindo números aos

objetos abarcados pela teoria, pode-se representar o seu peso. Além disso, se um conjunto de números pode fazer isso, outro conjunto de números relacionado ao primeiro também pode. No caso do peso, a multiplicação por uma constante pode transformar libras em quilos e vice-versa, sem que seja possível estabelecer qualquer parâmetro objetivo que permita escolher entre um e outro.

Com relação à teoria da semântica, segundo o autor, o conceito primitivo é o de *verdade*. A definição de verdade de Tarski modificada para se aplicar às linguagens naturais forneceria uma estrutura a partir da qual se poderiam descrever as condições de verdade de qualquer frase em uma língua específica, validando um manual de tradução. O manual de tradução de uma linguagem *A* para uma linguagem *B* correlacionaria as frases da linguagem *A* com as da linguagem *B*. Porém, há mais de uma maneira de se traçar essas relações. Diante de dois manuais de tradução diferentes para a mesma linguagem, considerando que há uma série de relações internas entre as frases desses manuais de tradução, não é possível retirar uma frase de um e simplesmente substituir por uma frase do outro, pois isso destruiria a trama daquele manual. O que é indeterminado, portanto, são os sistemas totais de tradução, isto é, os manuais inteiros. Assim como há indeterminação relativamente aos diferentes manuais de tradução, também são indeterminados os sistemas de atribuição de atitudes proposicionais, isto é, sistemas de interpretação do comportamento alheio.

Para Davidson, quando a indeterminação é aplicada às atribuições de atitudes proposicionais, funciona de modo análogo aos *sistemas de medição*. As medições da temperatura ou do peso, por exemplo, podem ser apresentadas de diferentes formas numéricas, mas há algo invariante sendo medido, aquilo que é realmente importante. Da mesma forma, as diferentes interpretações seriam diferentes versões para a mesma coisa, como medir a mesma coisa com os lados diferentes da trena.

A atribuição de atitudes é análoga, em muitos aspectos, à medição de várias magnitudes. Nós podemos atribuir números para captar os tamanhos, pesos e velocidades dos objetos, desde que os objetos exibam um padrão do tipo apropriado. Não supomos que haja entidades empíricas chamadas pesos, ou tamanhos ou velocidades as quais os objetos tenham. Como Carnap assinalou há muito tempo, nós não deveríamos pensar na frase 'Esta caixa pesa 8 libras' como identificando duas entidades denominadas 'o peso desta caixa' e '8 libras', mas sim como equacionando o peso da caixa em libras com o número 8. Logo, a ontologia exigida consiste dos objetos que têm pesos, e dos números. Os números não são parte dos objetos com peso; eles não pertencem ao mundo empírico, mas a nós, os que precisamos deles para captar certas relações entre os objetos. Da mesma forma, as entidades às quais nós relacionamos os pensantes quando nós atribuímos crenças e outras atitudes proposicionais a eles não estão nos pensantes - não nas suas mentes nem diante das suas mentes (DAVIDSON, 1997/2001h, p. 74-75, aspas internas do original, tradução nossa).

Em outro texto (1989/2001i), ele fala que quando atribuímos 345 gramas ao diamante de Koh-í-noor, não precisamos incluir gramas em nossa ontologia. Precisamos, sim, do objeto com peso que é o diamante. Gramas são unidades de medida que nos permitem representar as relações presentes nos objetos que têm peso. No mesmo sentido, as atribuições de atitudes proposicionais seriam formas de capturar os estados psicológicos de um sujeito. Não seria preciso, então, incluir em nossa ontologia¹¹³ as crenças atribuídas aos sujeitos como coisas distintas, mas precisamos do vocabulário psicológico para estarmos em relação com outros sujeitos e compreendermos seu comportamento.

Assim, o autor parece compreender a atribuição de estados mentais como uma espécie de artifício usado para interpretar o comportamento dos outros sujeitos. As atitudes proposicionais não são algo dentro da mente do sujeito, nem perante a mente que se possa agarrar por meio da atribuição de estados intencionais. Porém, o interessante é que *isso não leva Davidson ao antirrealismo*.

Os números podem captar pesos de objetos de infinitas maneiras, mas elas são conversíveis umas às outras, há algo de factual no peso dos objetos. Da mesma forma, há inúmeras maneiras de captar os estados mentais de um sujeito, o que de forma alguma vai implicar na ideia de que o sujeito não possua estados mentais. Note-se, porém, que admitir que o sujeito possui estados mentais não é admitir que esses estados mentais estejam fechados nele.

Eu penso que nem o indeterminismo mostra que as atitudes proposicionais sejam menos que plenamente reais (o que quer que isso possa significar) nem que nós devamos modificar o conceito de verdade quando falamos de atitudes proposicionais. Em outras palavras, muitas das nossas crenças e enunciados sobre o que as pessoas creem, pretendem, desejam e esperam são verdadeiras, e o são porque as pessoas têm essas atitudes (DAVIDSON, 1997/2001h, p. 112, tradução nossa).

4.1.3. Indeterminação da interpretação

Quando atribuímos atitudes proposicionais a outros sujeitos, essa atribuição pode ser verdadeira ou falsa acerca daquele sujeito. Quer dizer, podemos sempre nos enganar e atribuir a alguém crenças que não têm a ver com o que a pessoa pensa. Aqui, também, vale a analogia com os sistemas de medição. Eles não são entidades, mas a atribuição de um peso específico a um objeto pode ser verdadeira ou falsa.

¹¹³ Lembrando que, conforme o monismo anômalo, a ontologia inclui apenas uma realidade. Assim, não há entidades psicológicas independentes. O interpretivismo precisa de uma concepção de ontologia própria. Essa concepção se constrói a partir das relações entre o sujeito e o mundo físico e social.

Ao transportar para seu vocabulário o termo indeterminação da tradução alterando-o para *indeterminação da interpretação*, Davidson (1998/2005a, p. 316) acaba tendo que se defender da acusação de ter abandonado a tese de Quine. Ocorre que a indeterminação da interpretação é uma versão mais branda da indeterminação da tradução, pois Davidson vê algo fixo e estável sobre o qual ocorre a indeterminação. A interpretação, para ele, pode sofrer de indeterminação, mas a essa indeterminação não vaza para a realidade fixa das coisas.

Quando compara a atribuição de estados mentais aos sistemas de medição, Davidson defende que a indeterminação não é um golpe à realidade dos estados mentais, mas a mantém intacta. O realismo de Davidson se baseia em uma ligação entre crença e verdade. Considerando que a crença relaciona sujeito e mundo exterior, a verdade só pode ser abordada a partir das relações linguísticas que ocorrem de fato (DAVIDSON, 1990/2002a, p. 75). Ao mesmo tempo, essas relações linguísticas trazem sustentação à própria realidade dos estados mentais.

Ele rejeita o realismo em sua versão tradicional porque, ao estabelecer uma correspondência entre verdade e linguagem que independe da crença, a verdade acaba por se tornar incompreensível para os seres humanos. Também rejeita o antirrealismo, pois Davidson defende que existe um padrão para a verdade, e esse padrão é a intersubjetividade.

O realismo, com sua insistência sobre a correspondência radicalmente não-epistêmica, exige da verdade mais do que podemos compreender; o antirrealismo, com sua limitação da verdade ao que pode ser verificável, priva a verdade de seu papel enquanto um padrão intersubjetivo (DAVIDSON, 1990/2002a, p. 86)

Isso se mostra por meio da aprendizagem da linguagem. É preciso haver algo similar entre o aprendiz e o professor para que possa haver a aprendizagem de uma primeira língua: o mundo compartilhado e os comportamentos de assentimento e dissentimento perante as situações desse mundo. Esses comportamentos se tornarão progressivamente similares ao longo do processo de aprendizagem de uma língua. Isso também ocorrerá no processo de tradução de uma língua desconhecida.

A subdeterminação das teorias físicas é um importante argumento de Quine em favor da indeterminação da tradução, mas para Davidson há uma diferença entre a subdeterminação e a indeterminação. As teorias das ciências naturais falam sobre o que não foi observado, deixando-o subdeterminado em relação ao que é observado, mas nada impede que isso venha a ser determinado em algum momento no futuro. Já a indeterminação não é assim. “Nenhuma quantidade de evidência, finita ou infinita, poderia decidir se devemos medir áreas em acres ou hectares” (DAVIDSON, 1998/2005a, p. 318, tradução nossa). Nesse sentido, a

indeterminação acaba se revelando mais profunda e não deve ser confundida com a subdeterminação.

4.1.4. *A herança quineana e o caminho davidsoniano*

Conforme mencionado¹¹⁴, uma crítica apresentada por Davidson (1995/2005c) a Quine é que, de acordo com este, o que há de comum entre um falante de uma língua estrangeira e seu tradutor seriam os padrões de estímulos sensoriais. O problema, segundo Davidson, é que não temos nenhuma evidência disponível para sabermos os padrões de estímulos sensoriais ocorrendo nas outras pessoas. Os eventos externos ao falante e ao ouvinte, por outro lado, poderiam servir de base para uma interpretação das frases de observação, já que esses eventos estão disponíveis a ambos. A abordagem de Quine, para Davidson, é subjetivista demais.

Ao atribuir aos padrões de estímulos sensoriais o aspecto compartilhado entre o falante de um idioma nativo e seu tradutor, no modo de entender de Davidson, Quine constrói um esquema no qual o estrangeiro pode se enganar sistematicamente e não haverá jamais um parâmetro para que ele saia dessa situação de equívoco. Na concepção de Davidson, contudo, a indeterminação não precisa chegar a esse ponto, basta que se coloque o ponto compartilhado entre falante e intérprete não nos estímulos sensoriais, e sim nas situações observáveis compartilhadas entre eles, bem como em suas manifestações comportamentais frente às frases.

Davidson (1997/2001h, p. 77 e segs.) dedica-se a esclarecer os pontos de acordo e desacordo entre ele e Quine acerca da indeterminação. Ambos consideram que os comportamentos das pessoas perante as enunciações são as inferências básicas para a interpretação ou para a tradução. Por esse método, ambos consideram importante encontrar os padrões de assentimento e dissentimento com o objetivo de realizar a interpretação, e isso é feito com base nos fenômenos observáveis. Porém, diz Davidson (1997/2001h, p. 77, tradução nossa), “eu continuo, como Quine não faz, a usar o mesmo método para localizar os mecanismos para quantificação e referência cruzada. Portanto, ao contrário de Quine, eu não vejo a estrutura interna das frases mais simples como indeterminada”. Isso faz com que a compreensão que Davidson possui da indeterminação seja mais amena.

A indeterminação da interpretação, para Davidson, significa que “há mais de uma maneira de descrever o que é invariante” (DAVIDSON, 1998/2005a, p. 319, tradução nossa), ou seja, a realidade do mental não é negada por ele. Porém, isso que é invariante passará

¹¹⁴ Seção 2.1.

sempre pela interpretação, a qual parte do interlocutor, conforme as regras de racionalidade atribuídas por ele ao falante. Em outras palavras, por mais que se peça ao falante para esclarecer o seu discurso, esse esclarecimento sempre precisará passar por uma interpretação feita conforme as normas do ouvinte.

A linguagem que alguém fala é parte de todo um sistema de atitudes proposicionais. “Se nós mudamos nossa interpretação das palavras de uma pessoa, então, dada a mesma evidência total, nós devemos também mudar as crenças e desejos que atribuímos a ela” (DAVIDSON, 1979/1984g, p. 239, tradução nossa). Assim, a verdade das frases e a referência dos termos são ambas relativas à linguagem. Quando se muda a linguagem, uma série de ajustes serão feitos e é possível que as referências dos termos também mudem. Mas novamente, essa relativização da referência não implica em relativismo.

A concepção de indeterminação em Davidson é leve em comparação com Quine e, veremos em seguida, também em comparação com Dennett. Não há uma forma racional de escolher uma entre duas interpretações para o comportamento de um sujeito se ambas estiverem em harmonia com as evidências observáveis. Porém, essas duas interpretações não podem ser contraditórias, elas são apenas diferentes versões de uma mesma coisa.

4.1.5. A irredutibilidade tem a ver com a indeterminação?

Davidson (1998/2005a) reconhece que, em “Mental Events” (1970/1994b) havia atribuído a irredutibilidade do mental ao físico em parte à indeterminação da interpretação, mas cometera um erro. O erro está no fato de que a indeterminação se aplica amplamente sempre que houver a riqueza lexical necessária, e sempre há essa riqueza nas linguagens naturais. Porém, há algo fixo subjacente e há a possibilidade de conversão de uma forma à outra. “Tal indeterminação não ameaça a realidade do que é descrito” (DAVIDSON, 1998/2005a, p. 317, tradução nossa). Ou seja, na leitura de Davidson, há diferentes formas para apresentar a mesma realidade.

A indeterminação vai além da mera redundância das linguagens naturais, pois orações que sequer são sinônimas podem ser usadas para especificar o mesmo pensamento (1997/2001h, p. 77), mesmo assim, é um fato trivial. Os significados, segundo ele, podem ser capturados de diversas maneiras. Porém, há relações empíricas entre o falante, as suas frases e o seu ambiente que são fatos concretos, padrões invariantes. A indeterminação, então, não pode diferenciar o mental e o físico, pois ela está presente em qualquer discurso racional, tanto no discurso das ciências naturais quanto na psicologia.

A irredutibilidade se deve ao fato de que o discurso acerca do mental perderia o sentido se feito a partir de um vocabulário da física. A interpretação do comportamento de um sujeito exige caridade e racionalidade, por isso é algo extremamente complexo e traz uma forma de indeterminação mais profunda do que as medições dos objetos físicos. Essa é, inclusive, uma importante diferença entre os sistemas de interpretação e os sistemas de medição.

Porém, a irredutibilidade, além da indeterminação, também poderia à primeira vista ser considerada como um argumento pelo antirrealismo. A psicologia não pode ser acomodada por uma suposta teoria física capaz de unificar todas as coisas, porque as atitudes proposicionais têm seu conteúdo formado por intermédio de relações complexas entre sujeito, mundo e outros sujeitos ao longo de sua história de vida (DAVIDSON, 1997/2001h, p. 113). Isso não ocorre com os conceitos da física. Por isso, o vocabulário da psicologia não se encaixa ao da física. Davidson, contudo, não pensa que essa inadequação vocabular conduza ao antirrealismo acerca da psicologia.

Está claro para Davidson que a realidade das atitudes proposicionais não é atingida pela indeterminação. Ocorre que, conforme o monismo anômalo, os estados mentais são estados físicos descritos de outra maneira. Dependendo do que se busca explicar, o nível de descrição mais adequado será o mental ou o físico. Assim, se a ontologia das coisas físicas não está sendo posta em questão, também não há porque colocar em questão a ontologia dos estados mentais.

Na minha visão, o mental não é mais misterioso que a biologia molecular ou a cosmologia. Nossos conceitos mentais são essenciais para nossa compreensão do mundo como quaisquer outros; não conseguiríamos sem eles. As atitudes proposicionais, como intenções, desejos, crenças, esperanças, medos, são, cada pequeno pedaço, tão reais como átomos e tacos de beisebol, e os fatos sobre eles são tão reais como os fatos sobre qualquer outra coisa. Como poderia haver uma questão sobre a *ontologia* das entidades mentais para mim se, como sustentei, elas são idênticas a entidades que também descrevemos e explicamos, com termos diferentes, nas ciências naturais? (DAVIDSON, 1998/2005a, p. 316, grifo do original, tradução nossa)

É preciso encontrar uma justa medida para essa ontologia. Não é possível ser antirrealista sobre as atitudes proposicionais, pois a atribuição de atitudes proposicionais aos outros sujeitos é algo inevitável, não podemos ficar sem. Sem elas não poderíamos compreender nosso lugar no mundo físico nem social. Elas estão aí, só não devemos considerá-las como objetos fechados dentro nem fora dos sujeitos. “O único *objeto* que se requer para a existência de uma crença é um sujeito que creia [“a believer”]” (1997/2001h, p.

117, grifo do original, tradução nossa). Ter uma crença é estar em uma relação com uma proposição, não é como ter um objeto.

4.1.6. *Objetividade da subjetividade*

Ora, mas e quanto à autoridade de primeira pessoa? Aceitá-la não implicaria na negação da indeterminação da interpretação? Para Davidson, não. A resposta aqui é simples: a autoridade de primeira pessoa, segundo ele, restringe cada pessoa a apenas uma forma de interpretar a si mesma, porém, não restringe as interpretações que as outras pessoas terão. Em outras palavras, mesmo que o sujeito tenha apenas uma forma de apresentar seu pensamento, esse mesmo pensamento continua podendo ser apresentado também de outras formas (DAVIDSON, 1997/2001h, p. 124).

No fim das contas, a autoridade de primeira pessoa acaba sendo também um pilar que faz com que a atribuição de atitudes proposicionais seja tão objetiva a ponto de poder ser verdadeira ou falsa. Só é possível interpretar um sujeito atribuindo-lhe crenças e desejos com o pressuposto óbvio de que a pessoa é capaz de reconhecer suas próprias crenças e desejos. Conforme Silva Filho (2005, p. 165), “Para Davidson, a assunção da *autoridade da primeira pessoa* é uma condição de possibilidade para a atividade de interpretação” (grifo do original).

A objetividade das atitudes proposicionais se sustenta na intersubjetividade. Se o diálogo é possível, é porque (1) as atribuições de atitudes proposicionais tendem à verdade; (2) as nossas próprias crenças tendem a ser formadas a partir de relações causais com o mundo objetivo ao longo de nossa história de vida e (3) nossas demais crenças (além das atribuições de atitudes proposicionais) são majoritariamente verdadeiras. Até mesmo o mais delirante dos lunáticos, segundo essa concepção, terá esse fundo de crenças majoritariamente verdadeiras.

Para que se possa trazer algum fundamento à verdade e à objetividade, é preciso negar o subjetivismo. Uma visão de mundo subjetivista acerca do mental leva ao ceticismo acerca do mundo exterior e acerca das outras mentes. De acordo com essa visão, entre mente subjetiva e mundo, haveria esquemas conceituais capazes de permitir à mente a interpretação desse mundo. Davidson rejeita esses esquemas. A relação entre mente e mundo passa a ser direta. Assim, ao se compreender o mental de modo conectado tanto ao mundo exterior quanto às outras mentes, essas formas de ceticismo deixam de fazer sentido.

A negação do dualismo esquema/conteúdo como o terceiro dogma do empirismo tem como consequência a não aceitação de uma posição relativista. As diferentes interpretações

orbitam uma base fixa. Nada mais são do que formas diferentes de dizer a mesma coisa. Conforme El-Hani e Pihlström (2005, p. 233):

[...] Davidson, identificando esquemas conceituais com linguagens, parte da premissa de que a tradutibilidade para a *nossa* linguagem é um critério necessário para que sejamos capazes de entender qualquer coisa (qualquer esquema conceitual possível) como linguagem (ou um esquema conceitual); portanto, argumenta ele, a noção de esquemas radicalmente diferentes, não traduzíveis ou incomensuráveis, não faz sentido. Se algo fosse *realmente* incomensurável com, ou radicalmente diferente do, “nosso esquema” (qualquer que seja ele) ou nossa linguagem, não teríamos qualquer maneira de considerar aquele algo como um fragmento de linguagem, ou um esquema (ou até mesmo como comunicação inteligível), em primeiro lugar (grifos do original).

A noção de intérprete radical vem nos ajudar a compreender isso. Para Davidson, não há esquemas conceituais que sejam intraduzíveis um ao outro e que sirvam de intermediário entre sujeito e realidade objetiva. Por isso, o intérprete radical é capaz de se inserir em um processo de triangulação para poder compreender um discurso em qualquer língua totalmente desconhecida. Esse processo não tem intermediários. Ele se constitui de relações diretas entre mundo e sujeitos em comunicação entre si. A triangulação acaba por trazer ao discurso um aspecto fixo que subjaz aos diferentes idiomas: uma espécie de linguagem do intérprete radical, a linguagem na qual ele compreende as outras linguagens.

Do ponto de vista de um intérprete comum, não é possível nem a atribuição de crenças majoritariamente falsas nem um desacordo massivo entre ele e o sujeito da interpretação. Para começar, o intérprete precisa presumir a ampla verdade do corpo de crenças daquele sujeito. Mas isso se revela mais do que uma presunção graças à ancoragem na realidade objetiva dada pela possibilidade de comunicação, resultado da interpretação bem sucedida.

Como apontado por Stroud (2005), os argumentos de Davidson não impedem, a princípio, a possibilidade de um conjunto de crenças ser todo ou majoritariamente falso, desde que essas crenças não se sujeitem ao processo de interpretação. Se houver uma interpretação bem sucedida de um sujeito, pode-se concluir que suas crenças são majoritariamente verdadeiras. Isso não significa que existam, na realidade, sujeitos ininterpretáveis. Essa não é a tese de Stroud, muito menos a de Davidson. Significa que, em tese, poderia haver. Por isso, embora o argumento do intérprete radical impeça a formulação do ceticismo global, ele não é capaz de refutá-lo.

Manter o realismo acerca das atitudes proposicionais é uma liga importante no pensamento de Davidson como um todo. Se o autor partisse para o antirrealismo, o argumento

da triangulação perderia sua força contra o ceticismo¹¹⁵, e o monismo anômalo não teria mais qualquer sentido, tornando-se vazio. Além disso, o princípio de caridade flutuaria como um barco à deriva, sem nada em que se ancorar. Uma vez que o pensamento de Davidson inclui filosofia da linguagem, filosofia da mente, epistemologia e ontologia, a atribuição de um status ontológico às atitudes proposicionais acaba sendo um ponto de contato entre todos esses ramos do pensamento dele.

4.2. Dennett: a realidade que provém da utilidade¹¹⁶

De acordo com Dennett (1991a, p.27), em geral há apenas duas posições acerca da realidade das crenças, ambas extremadas: ou os filósofos adotam um realismo forte ou um materialismo eliminativo, mas nenhuma dessas duas posições, para ele, seria a mais apropriada para lidar com a questão. Uma metáfora que perpassa vários dos textos de Dennett é a comparação entre os centros de gravidade e as crenças (ao longo de sua obra, os centros de gravidade se delineiam como uma metáfora não apenas para as crenças, mas para o conjunto das crenças e para a própria consciência). Centros de gravidade são apenas ficções úteis para os cientistas explicarem certos aspectos da realidade, ou eles são perfeitamente reais?

O ponto de equilíbrio de todas as forças gravitacionais de um corpo não possui uma existência concreta. Centros de gravidade são objetos abstratos definíveis por meio de forças físicas e outras propriedades. Dennett (1992a) apresenta diversas características dos centros de gravidade as quais o ajudam a estabelecer seu status. Segundo ele, os centros de gravidade não são itens físicos do mundo, embora sejam um conceito da física como disciplina. Suas únicas características físicas são localização espaço-temporal. (Importante notar que, embora haja várias analogias, esta é uma desanalogia entre os centros de gravidade e os estados mentais, pois não conseguimos definir a localização espaço-temporal dos estados mentais, sequer podemos pensar que eles estejam em algum lugar no corpo de um sujeito, uma vez que se determinam nas relações daquele sujeito com o mundo externo e demais sujeitos.) Eles se movimentam, mas de modo descontínuo, de modo que têm uma história. Não possuem textura nem cor. São objetos puramente abstratos, ficções teóricas com um papel muito bem

¹¹⁵ De acordo com Pereira (2016), embora Davidson esteja ciente de que o argumento da triangulação bloqueia a postulação dos ceticismos acerca das outras mentes e acerca do mundo exterior, não era esse seu objetivo principal. Ele estava mais preocupado em rejeitar a noção de mental como algo intrinsecamente subjetivo. O ofuscamento do ceticismo é uma consequência. De todo modo, vale notar que essa consequência acaba por reforçar o argumento.

¹¹⁶ Apresentado ao grupo MELE/Fil-UnB (Mente, Linguagem e Evolução) em dez./2017. Agradeço aos membros do grupo pelos comentários, em especial, ao professor Paulo Abrantes pelas sugestões de leitura.

delimitado na física. Cada objeto concreto possui um centro de gravidade o qual pode ser manipulado por nós.

Um centro de gravidade é apenas um *abstractum*. É apenas um objeto ficcional. Mas quando eu digo que ele é um objeto ficcional, não quero desacreditá-lo, é um objeto ficcional maravilhoso e tem um lugar perfeitamente legítimo dentro da ciência física séria, sóbria, genuína (DENNETT, 1991a, p. 276, tradução nossa).

Para Dennett, mais importante que o status ontológico das atitudes proposicionais é a sua *utilidade*. A realidade, para ele, provém da sua força epistemológica, ou seja, atitudes proposicionais são reais na medida em que são fundamentais para a nossa compreensão do mundo que nos cerca. Assim como os centros de gravidade, as atitudes proposicionais são tomadas por Dennett como objetos abstratos, não têm localização concreta, não podem ser tocados, mas desempenham um papel central na nossa compreensão da realidade circundante. Sem as atitudes proposicionais, não poderíamos nos comunicar com os outros. Elas são imprescindíveis para que possamos dar algum sentido à realidade e planejar nossos próximos passos nas relações com os outros.

Dennett argumenta que “a questão de se os objetos abstratos são reais - a questão de se ‘se deve ser realista sobre eles’ ou não - pode tomar dois diferentes caminhos, os quais nós devemos chamar de o metafísico e o científico” (1991a, p. 28, aspas duplas do original substituídas por aspas simples, tradução nossa). O metafísico aborda a realidade ou existência dos objetos abstratos sem os diferenciar em termos de sua utilidade científica. O problema é que há diferentes tipos de objetos abstratos, e seu status não parece ser o mesmo, ainda que sejam semelhantes em seu aspecto metafísico. Por exemplo, algumas pessoas relatam ouvir vozes misteriosas que conversam com elas, podendo assombrá-las ou mesmo ajudá-las. Mas essas vozes não parecem ter o mesmo status dos centros de gravidade, embora ambas possam ser consideradas objetos abstratos. São muito diferentes! Nesse sentido, então, o caminho da utilidade é considerado pelo autor um ponto de partida bastante relevante para se pensar acerca do realismo. Quanto aos centros de gravidade:

Centros de gravidade são reais porque são (de algum modo) *bons* objetos abstratos. Eles merecem ser levados a sério, aprendidos, usados. Se formos longe o suficiente para distingui-los como *reais* (contrastando-os, talvez, com aqueles objetos abstratos que são *falsos*), é porque nós pensamos que eles servem como representações claras de forças reais, propriedades “naturais” e coisas do gênero (DENNETT, 1991a, p.29, grifos e aspas do original, tradução nossa).

É possível, por exemplo, definir com precisão o ponto central a toda a população brasileira, embora esse ponto mude todo o tempo. Outro exemplo do autor é o centro das meias desaparecidas de Dennett, um ponto imaginário localizado exatamente ao centro de um

círculo traçado ao redor de todas as meias que Dennett perdeu ao longo de sua vida (1991a, p.28). Seria bastante útil saber a localização de algumas dessas meias, mas saber o centro das meias desaparecidas não tem qualquer utilidade, embora seja um objeto abstrato muito específico e bem definido, parecendo ter um status ontológico perfeitamente real. Todos esses centros - o centro da população brasileira, o centro das meias desaparecidas de Dennett e os centros de gravidade - possuem exatamente o mesmo status ontológico, são objetos reais, precisos, bem definidos, mas no fim das contas isso diz muito pouco sobre eles. Nesse sentido, Dennett defende que o status dos objetos abstratos não seja estudado por essa perspectiva metafísica ampla, mas sim pela perspectiva de sua utilidade. No caso das atitudes proposicionais, estamos diante de algo mais do que útil, algo imprescindível do ponto de vista epistemológico.

Os objetos abstratos que tendem a ser reais, para ele, são aqueles que nos permitem uma melhor relação epistêmica com o mundo. Os centros de gravidade são bons objetos abstratos porque são úteis na compreensão dos objetos físicos. Dennett considera a perspectiva científica como uma alternativa viável para responder à pergunta acerca da realidade dos centros de gravidade: o caminho da ciência, de acordo com ele, é o da utilidade¹¹⁷. O caminho metafísico, se for isolado das considerações acerca dos resultados práticos de determinada consideração teórica, não será capaz de produzir resultados frutíferos. Para Dennett, no geral, a filosofia se torna muito estéril se ficar fechada em si mesma. É preciso dialogar com as práticas quotidianas e as científicas.

O status das crenças, para ele, será melhor abordado se for considerado equivalente ao dos centros de gravidade, isto é, objetos abstratos que provavelmente são reais. Essa realidade provém do fato de que são extremamente poderosos na nossa compreensão da realidade. O autor busca um tipo de *realismo intermediário*. As crenças não são apenas ficções, mas também não têm o status ontológico de um objeto perfeitamente real dentro do sujeito ou um estado físico correspondente no cérebro (DENNETT, 1987, 14). O ponto chave a ser

¹¹⁷ Isso poderia dar muita discussão sobre realismo e instrumentalismo em filosofia da ciência. Sem entrar nesse debate, podemos apenas pensar que, dentro de uma comunidade científica, há vários fatores que a levam à tomada de decisões teóricas e a utilidade é um fator importante. Aquela decisão ajuda a compreender aspectos da realidade que antes pareciam nebulosos? Ela é coerente com outras crenças valiosas para aquela comunidade científica? Porém, após adotados certos conceitos, independente das razões pelas quais eles tenham sido adotados, os cientistas aparentemente se apegam a eles como aspectos da própria realidade. Faz sentido passar do útil para o real? Ora, se algo explica tão bem tantas coisas que pareciam misteriosas, que razões teríamos para continuar considerando-o como apenas útil? Isso pode se aplicar à atribuição de crenças. Se conseguimos explicar e prever tão bem o comportamento de nossos pares por meio da atribuição de crenças, então porque haveríamos de considerar isso como apenas uma atribuição instrumental sem lastro na realidade?

considerado para a compreensão do status das crenças é a psicologia de senso comum¹¹⁸. Ela é uma habilidade natural sem a qual as pessoas não poderiam se relacionar.

4.2.1. Padrões de crenças

Na concepção de Dennett (1991a), a capacidade de interpretar as ações alheias depende da capacidade de predizê-las. Mas a previsão é uma capacidade que possui raízes indutivas, e portanto depende da percepção de *padrões*, sequências regulares e discerníveis. Nós prevemos o comportamento dos outros criando expectativas de padrões. Por isso, Dennett considera essa noção altamente relevante para lidar com a questão da atribuição de estados mentais. O que se busca compreender, agora, não é mais o status das crenças, e sim o status dos padrões de crenças, pois é ele que importa para as atividades e planos que realizamos em nosso dia a dia. Nesse sentido, ele comparou as posições de alguns autores acerca da realidade das crenças, situando sua própria posição entre eles, como uma posição intermediária ao realismo e ao instrumentalismo. Os autores comparados foram Fodor, os quineanos (isto é, o próprio Dennett e Davidson), Rorty e Churchland.

A posição de Fodor é de *realismo intencional forte*. Conforme essa concepção, o padrão de crenças de um sujeito seria um padrão de estruturas no cérebro, caso contrário, não poderia ser um padrão discernível. A posição dos quineanos, entre os quais Dennett e Davidson se situam é de que o padrão de crenças estaria no comportamento dos agentes - no caso de Davidson, agentes sujeitos à interpretação radical e, no caso de Dennett, à postura intencional. Essa posição seria denominada *realismo de força regular*. Porém, Dennett considera a força de seu realismo menor do que a de Davidson, denominando a própria posição de *realismo brando* ["mild realism"]. A posição de Rorty é chamada de *irrealismo brando*, segundo o qual o padrão de crenças estaria apenas nos olhos dos expectadores, uma posição mais instrumentalista. Por último, a mais irrealista das posições, o *materialismo eliminativo* de Churchland que nega a realidade das atitudes proposicionais. Ora, a posição mais apropriada, segundo Dennett, é o realismo brando acerca dos padrões de crenças distinguíveis por meio da postura intencional¹¹⁹.

¹¹⁸ Ver: seções 3.2 e 3.2.1.

¹¹⁹ Conforme Abrantes (2013), Sterelny e Godfrey-Smith são realistas acerca da atribuição de crenças e desejos apenas no que concerne às mentes humanas. Em relação aos animais não humanos, por outro lado, não seria apropriado considerá-los capazes de atribuição de estados mentais.

Essas diferenças não são acerca da realidade última das coisas físicas ou entidades abstratas. São diferenças de opinião acerca de até que ponto os estados mentais em geral, incluindo as atitudes proposicionais, são reais como os átomos ou os centros de gravidade.

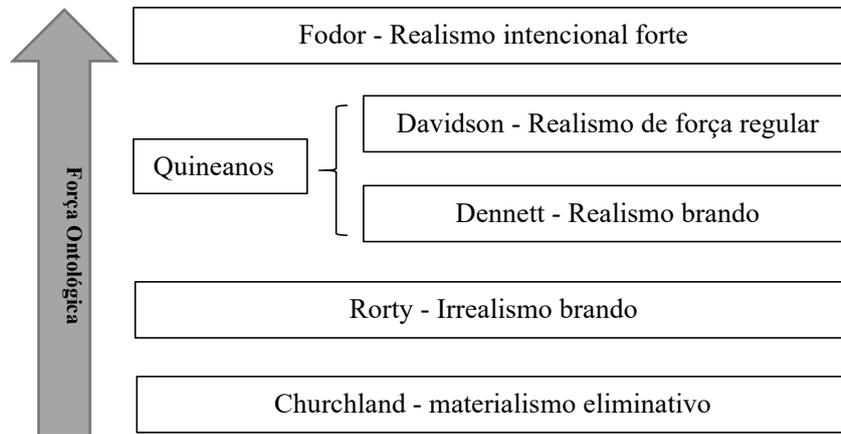


Figura 1: Posições sobre a ontologia dos padrões de crenças

Dennett considera sua posição como sendo intermediária entre o realismo intencional forte de Fodor e o instrumentalismo. Os padrões são importantes para sua posição, pois, segundo ele, os padrões são, por definição, *discerníveis*. O que é discernível, é discernível por alguém, sob alguma perspectiva (o que não significa que todas as pessoas sejam capazes de discernir). Assim, para a posição de Dennett, focada no observador e na sua perspectiva, a concepção de padrão ganha um papel esclarecedor na compreensão do status das atitudes proposicionais.

Um animal com aparato sensorial diferente perceberá padrões na natureza que são imperceptíveis para nós, ou seja, *os padrões dependem do ponto de vista*. Mas isso também acontece entre as pessoas, padrões diferentes são percebidos, o que não indica um desacordo (seria talvez um desacordo uma pessoa perceber um padrão onde outra não percebe nenhum). As capacidades de perceber padrões variam, inclusive entre diversos momentos da história de vida de uma mesma pessoa. O tempo e os recursos disponíveis para a identificação de padrões também são fatores capazes de alterar os padrões percebidos.

Dennett enfatiza o fato de que nossos aparatos cognitivos são bastante voltados para a percepção dos padrões na natureza, uma vez que isso permite o ajuste de nossas decisões comportamentais. Outro ponto importante: embora os padrões sejam algo discernível por alguma criatura, isso não significa que eles estejam na criatura que percebe (1991a, p.37). O padrão é *real*, ele está nas coisas na medida em que é possível descrevê-lo e serve como base

para realização de previsões. As criaturas percebem esses padrões conforme suas habilidades, tanto as inatas quanto as aprendidas ao longo da vida. Por isso, duas pessoas que percebam padrões diferentes diante dos mesmos dados poderão fazer previsões iguais, mas também poderão fazer previsões diferentes, o que não faz com que o padrão seja mais ou menos real. A percepção do padrão influencia na taxa de sucesso das previsões, mas está sujeita a erros e ruídos.

Por exemplo, diante de uma sequência de dados que apresente um padrão visual, três pessoas podem ser chamadas a fazer previsões acerca do que se seguirá dali. A primeira pessoa pode bater o olho, encontrar um padrão bem simples e fazer uma previsão rápida, mas com alta taxa de ruído. A segunda pessoa poderá analisar um pouco mais detidamente os dados, realizar alguns cálculos mentais e, em seguida, fazer uma previsão mais acurada. Já a terceira pessoa poderá lançar mão de instrumentos de medição e uma máquina calculadora para encontrar padrões mais precisos e complexos e demorar um pouco mais, encontrando um padrão que lhe permita uma taxa bastante alta de sucesso em suas previsões.

O que se deve notar aqui é que, embora tenham diferentes taxas de ruído, os três padrões estão presentes, são *padrões reais* naquela sequência de dados. Qual é o melhor? Isso dependerá da situação. Às vezes é preciso fazer previsões rápidas e econômicas, ainda que pouco precisas. Em outros momentos, contudo, é preciso haver alta taxa de precisão, com investimento alto de tempo e energia, pois seria arriscado demais lidar com a taxa de ruído. Aqui, mais uma vez, Dennett enfatiza o status de utilidade como anterior ao status ontológico.

Prever que alguém se esquivará se você lhe atirar um tijolo é fácil a partir da psicologia de senso comum, mas isso será sempre intratável se você tiver que traçar os fótons do tijolo até o globo ocular, os neurotransmissores do nervo ótico ao nervo motor e daí em diante (DENNETT, 1991a, p. 42, tradução nossa).

A adoção da postura intencional nos permite a identificação de certos padrões que não poderíamos identificar por meio da postura física nem pela postura de projeto. Ela realiza previsões altamente confiáveis do comportamento humano e não humano. Os diversos autores tratados acima concordam em relação à importância da psicologia de senso comum, mas divergem acerca das implicações ontológicas disso para a realidade das atitudes proposicionais. Porém, um ponto importante a se notar é que, por mais reais que sejam os padrões, isso não pressupõe a existência de algo correspondente na estrutura cerebral, como quereria Fodor.

4.2.2. *Dennett critica Davidson acerca da ontologia das atitudes proposicionais*

Dennett dedica parte do artigo *Real Patterns* (1991a) à apresentação da diferença - que considera sutil - entre sua perspectiva e a de Davidson acerca da ontologia das atitudes proposicionais. Ambos pensam que as atitudes proposicionais são reais a ponto de permitirem uma previsão acurada do comportamento alheio, mas Davidson tem uma postura mais realista. Considerando que ambos foram discípulos de Quine, Dennett vê como a raiz dessa diferença opiniões diversas sobre o status do princípio de indeterminação da tradução, embora o princípio seja aceito pelos dois autores. Conforme Miguens (2002, p. 172) “Para Dennett, Davidson ainda que considere acertadamente que a teoria do mental deve ser uma teoria da interpretação, leva as atitudes proposicionais mais a sério do que deveria, como quineano. Dennett reclama para si, por conseguinte, o título de verdadeiro quineano.”

Segundo Dennett (1991a, p. 46), Davidson vê esse princípio de modo quase trivial. Ele encara dois manuais de tradução que sejam diferentes e igualmente funcionais analogamente a duas escalas diferentes para medir algo, por exemplo, a medição da temperatura em graus fahrenheit ou célsius. Foi dito acima que, assim como os números podem capturar todas as relações empiricamente significativas entre pesos e temperaturas de infinitas formas possíveis, Davidson defende que a fala de uma pessoa poderia capturar todas as características significativas do pensamento de outra pessoa, sem que isso desafie a realidade das atitudes proposicionais.

Dennett, por outro lado, vê o princípio de indeterminação de outra maneira. Segundo ele, diante de duas interpretações rivais do comportamento de um indivíduo, o que estaria acontecendo é que aqueles dois intérpretes apresentam uma concordância na forma geral do conjunto de crenças atribuído. Porém, não concordam ponto a ponto porque percebem padrões diferentes - embora não discordantes - no comportamento daquele indivíduo. Para Dennett, o ponto principal da tese da indeterminação é que a interpretação não precisa ter um lastro subjacente.

A diferença, portanto, é que, em Davidson, parece haver algo fixo escondido sob as diferentes atribuições de padrões de crenças e desejos, ao passo que Dennett não está comprometido com essa ideia, e por isso rejeita a analogia dos diferentes sistemas de medição. O pensamento de Davidson acaba tendendo mais ao realismo do que o de Dennett porque a indeterminação, para o primeiro, consiste apenas em diferentes olhares sobre a mesma realidade. Já para o segundo, os olhares diferentes, por definirem padrões diferentes, descrevem também a realidade de maneiras diferentes, já que os padrões estão na realidade.

Isso tem como consequência que, pela perspectiva de Dennett, mesmo que dois sistemas de atribuição de atitudes proposicionais façam previsões diversas, se ocorrer de apenas uma se realizar, isso não definiria qual deles seria o correto, pois a realização da previsão pode ter simplesmente caído na taxa de ruído.

Como isso me faz menos realista do que Davidson? Eu vejo que poderia haver dois sistemas diferentes de atribuição de crença a um indivíduo que diferissem *substancialmente* no que eles atribuísem - mesmo permitindo predições substancialmente diferentes do comportamento futuro do indivíduo - e ainda não haveria um fato mais profundo que pudesse estabelecer que um fosse uma descrição das crenças *reais* do indivíduo e o outro não. Em outras palavras, poderia haver dois padrões diferentes, mas igualmente reais discerníveis no mundo ruidoso. Os teóricos rivais não precisariam nem concordar sobre que partes do mundo seriam padrões e quais seriam ruído, e mesmo assim nada mais profundo iria resolver o problema. A escolha de um padrão iria certamente ser afinal do observador, um assunto a ser decidido sobre um solo pragmático idiossincrático. Eu mesmo não vejo qualquer característica na posição de Davidson que pudesse ser um obstáculo sério para suas analogias astuciosas e concordância comigo. Mas então ele precisaria admitir que a indeterminação não é um assunto tão trivial no fim das contas (DENNETT, 1991a, p. 49, grifos do origina, tradução nossa).

Essa diferença entre Davidson e Dennett acerca da ontologia das atitudes proposicionais não deixa de ser sutil. Os dois autores reconhecem que somos todos fiéis à psicologia de senso comum, confiamos nela e baseamos nela as nossas tomadas de decisão cotidianas. Trabalham sobre isso, reconhecendo a atribuição de crenças como o aspecto mais relevante na compreensão do mental. A psicologia de senso comum é uma ferramenta preditiva que, para eles, descreve padrões reais, ainda que haja divergência acerca da sua força ontológica.

A posição intermediária de Davidson, assim como a minha, vincula a realidade à existência bruta do padrão, mas Davidson negligenciou a possibilidade de dois ou mais padrões *conflitantes* serem sobrepostos aos mesmos dados - uma indeterminação da tradução mais radical do que ele supôs possível (DENNETT, 1991a, p. 51, grifo do origina, tradução nossa).

Aparentemente, isso não foi mera negligência de Davidson. Ele está empenhado em defender a existência de algo fixo circundado pelas interpretações variantes. Para Davidson, há algo fixo por trás dos diferentes padrões de atribuições de estados mentais, vale enfatizar. Por isso ele compara as diferentes atribuições de atitudes proposicionais a diferentes sistemas de medição. Também por isso ele proíbe que as interpretações variantes sejam logicamente incompatíveis entre si.

Por exemplo, Paul Churchland (1984/2004) defendeu que, embora a psicologia de senso comum tenha um papel importante, a neurologia estaria em vias de desenvolver um conjunto de padrões com poderes preditivos superiores. Dennett, por sua vez, não considera

esse padrão da neurologia atrativo por ser dispendioso demais para nossas necessidades cotidianas. Em outras palavras, ainda que tenhamos em nossas mãos um conjunto de padrões altamente poderoso proposto pela neurologia para nossas explicações e previsões do comportamento, vamos continuar utilizando a psicologia de senso comum, pois ela faz parte daquilo que nos constitui como humanos. Por mais que a neurologia se desenvolva na explicação do comportamento, portanto, ela continuará sendo irrelevante em nossas vivências diárias.

Fodor, com seu realismo intencional fortíssimo e Churchland, com seu materialismo eliminativo são representantes das posições extremas. Porém, essas posições acabam por subestimar, cada uma à sua maneira, a realidade da psicologia de senso comum. Davidson e Dennett se encontram em posições intermediárias. Quanto à pergunta sobre se sua posição é instrumentalista ou realista, Dennett prefere não apresentar uma resposta. Contudo, diante da ênfase que Dennett dá à utilidade quando trata do status da atribuição de estados mentais, podemos considerar sua posição um interpretivismo que não é genuinamente instrumentalista, pois confere força ontológica às atribuições de crenças, mas tem um aspecto instrumental. Podemos entender dessa forma porque, para Dennett, as atribuições de estados mentais são um *instrumento* de vital importância para nós, donde provém sua *realidade*. Quanto à posição de Davidson, a força ontológica das atribuições de estados mentais é ainda um pouco maior. Ele tende ao realismo cotidiano¹²⁰.

4.2.3. Resposta de Davidson a Dennett

Nas perspectivas desses autores, Dennett e Davidson, quando atribuímos conjuntos de atitudes proposicionais a um sujeito, há uma tendência forte a que (1) muitas daquelas atitudes proposicionais atribuídas realmente pertençam ao sujeito e (2) que o sujeito esteja majoritariamente justificado em ter aquelas atitudes proposicionais. Isso ocorre, em Davidson, pela forma como aprendemos a linguagem, ou seja, o processo de triangulação relaciona nossas crenças, as crenças do nosso interlocutor e o mundo. Por mais que as crenças se justifiquem umas pelas outras, há algum tipo de contato com o mundo que permite a comunicação. No caso de Dennett, a questão se torna mais sutil. Não há nada tão firme ancorando as atitudes proposicionais, mas nossa espécie chegou até aqui atribuindo crenças e desejos aos outros¹²¹. Nossas atribuições de atitudes proposicionais cotidianas são, além de

¹²⁰ Ver: seção 2.2.1.

¹²¹ Adiante, será possível perceber que o processo evolutivo acaba por trazer alguma firmeza às atribuições de

úteis, vitais.

Um ponto comum entre esses autores que se torna importante para compreendermos a sutileza de seu realismo é que ambos defendem um tipo de monismo não reducionista. Do ponto de vista ontológico, há apenas uma realidade. Mas os diversos tipos de situações dessa realidade exigem vocabulários específicos. Então, o ponto a ser destacado aqui é que, para eles, o vocabulário da psicologia não poderá ser reduzido ao da física. Tal redução implicaria em perdas epistemológicas que vão além do que poderíamos humanamente suportar.

Não poderíamos interagir com os outros sem lhes atribuir estados psicológicos. Isso traz um realismo às atitudes proposicionais do qual não é possível abrir mão. Sua realidade, no caso de Dennett, está atrelada à sua força epistêmica. No caso de Davidson, o realismo das atitudes proposicionais está vinculado também ao argumento da triangulação, um argumento que afasta o ceticismo do caminho, dando consistência às crenças, às atribuições de crenças e ao mundo objetivo.

Convém notar que, quando falamos de verdade em Davidson, há a verdade objetiva, independente das crenças, embora não haja a maneira certa de descrevê-la, várias descrições diferentes da mesma realidade são igualmente válidas. Mas e quanto à atribuição de atitudes proposicionais? Acontece o mesmo, pois as atribuições de atitudes proposicionais podem ser verdadeiras ou falsas, ainda que não haja a forma unívoca de interpretar o comportamento e o discurso de alguém. Diferentes interpretações do comportamento de um agente, para ele, preservam o sentido geral das atitudes proposicionais dele, de modo que, em princípio, seria possível confrontar a interpretação e as crenças do agente para dizer se a interpretação é correta.

Dennett (1991a) considera sua perspectiva acerca da ontologia das atitudes proposicionais próxima da de Davidson. Porém, Dennett pensa que Davidson passa do ponto no realismo ao comparar as atitudes proposicionais aos sistemas de medição, pois diferentes sistemas de medição, em geral, se sobrepõem a algo fixo que está subjacente. De fato, Davidson adota uma postura mais branda que a de Dennett em relação à indeterminação - o que acaba intensificando seu realismo - e procura defender seu ponto. Acerca de Dennett, escreve Davidson:

[E]le pensa que dois diferentes sistemas de atribuição de crenças a um indivíduo podem diferir substancialmente, até o ponto de produzir diferentes previsões do comportamento, e mesmo assim nada iria estabelecer que um sistema e não o outro descreveria as crenças reais de uma pessoa (Davidson, 1997/2001h, p. 81, tradução nossa).

De fato, Dennett (1991a) faz uma leitura mais radical da indeterminação e, provavelmente, mais próxima do próprio Quine. Davidson reconhece isso ao apresentar a comparação com os sistemas de medição. “[A] indeterminação da interpretação é a contraparte semântica da indeterminação da tradução de Quine. Em minha abordagem, o grau de indeterminação será, eu penso, menos do que Quine contempla” (DAVIDSON, 1975/1984d, tradução nossa).

Davidson (1995/2005c) está preocupado com mostrar que a indeterminação não implica em relativismo acerca da verdade. Para ele, duas teorias diferentes não podem ser “logicamente incompatíveis e empiricamente equivalentes” (1995/2005c, p. 76, tradução nossa), como quereria Quine (1970). As teorias diferentes seriam descrições da mesma realidade em termos diferentes. Elas não seriam contraditórias entre si, mas conteriam ambiguidades semânticas relativamente a termos-chave, o que lhes daria essa aparência de contradição. Ou seja, a partir de um olhar bem aproximado para essas teorias, perceber-se-á que, no fundo, elas falam idiomas diferentes. Essa manobra em relação à indeterminação, para Davidson, é importante para evitar a posição relativista segundo a qual dois conjuntos de crenças incompatíveis poderiam ser ambos verdadeiros, pois isso desconectaria a verdade da realidade objetiva.

Conforme Pereira (2016), embora Davidson tenha sempre tomado a verdade como indefinível, ele tentou em alguns momentos esclarecê-la. Inicialmente, ele a considera como correspondência. Porém, posteriormente, passa a considerar que a coerência produz correspondência. As crenças se justificam umas sobre as outras, por coerência, mas a existência de comunicação entre as pessoas é evidência da existência do tripé que conecta os sujeitos em comunicação com o mundo objetivo. Assim, é o mundo que causa parte de nossas crenças, as quais constituirão uma rede de crenças majoritariamente coerentes entre si.

Silva Filho (2002, p. 154) argumenta que, embora Davidson não traga uma resposta acerca do que seja uma crença verdadeira (até porque não há crenças isoladas que possam se confrontar com a realidade), ele consegue mostrar “porque é que as crenças verdadeiras são justificadas”. Estamos justificados, a partir da coerência entre nossas crenças, em supor que elas são majoritariamente verdadeiras.

O conceito de verdade, assim como o de objetividade, está sempre ligado ao surgimento do pensamento e da linguagem. Na concepção de Davidson, a noção de objetividade só passa a ter algum sentido a partir da comunicação com os outros. A comunicação faz com que passemos a diferenciar o que se passa nas mentes dos outros, nas nossas mentes e no mundo. Isso não significa que não haja um mundo independente das

mentes. Há a verdade objetiva e independente das mentes, mas ela não teria qualquer *sentido* se não houvesse mentes. É assim que a verdade se torna um elo entre eu, o mundo e o outro sujeito. “[O]s três conhecimentos - sobre o mundo, nossas mentes e as mentes dos outros - são assimétricos, mas se referem à mesma realidade (SILVA FILHO, 2005, p. 166)”.

O princípio de caridade tem um papel importante no realismo de Davidson, pois em última instância, ele concatena os estados mentais dos sujeitos em comunicação com o mundo, fazendo com que a atribuição de estados mentais se fundamente na intersubjetividade. Quando duas pessoas se comportam de modo a assentir às mesmas frases pelos mesmos eventos ou situações, na concepção do autor, elas compartilham de muitas frases de observação.

Verdade e significado andam juntos no pensamento de Davidson. “O centro da vida do significado passou a coincidir com o centro de gravidade que determina a verdade” (DAVIDSON, 1995/2005c, p.80, tradução nossa). Davidson não desenvolve a noção de centro de gravidade, como faz Dennett, mas é interessante a sua colocação da verdade como um centro de gravidade. Há uma forte tendência do conjunto de nossas crenças a serem majoritariamente verdadeiras, embora tenhamos uma série de crenças falsas. Nosso conjunto total de crenças orbitam ao redor desse centro de gravidade. Por mais difícil que seja pinçar as crenças verdadeiras e separá-las das falsas, a verdade predomina como um centro de atração.

4.2.4. *Moedas como sistemas de medição*

Tomem-se os sistemas monetários como sistemas de medição. Foi dito que Davidson compara a noção de indeterminação a sistemas de medição. Afinal, dois sistemas de medição podem ser convertidos um ao outro e medem as mesmas características. Há uma proporcionalidade entre eles que costuma ser baseada em aspectos estáveis da realidade. Podemos, por exemplo, converter facilmente as polegadas a centímetros, isso não fará qualquer diferença em relação ao que está sendo medido. Há uma indeterminação, pois não existem critérios racionais que nos permitam priorizar um sistema de medição em detrimento do outro, mas há algo fixo subjacente ao que está sendo medido.

Contudo, cabe aqui lembrar dessa espécie de sistema de medição onde as diferentes unidades são conversíveis umas às outras, mas o que medem é altamente difuso e instável: as moedas. Podemos converter reais em euros, pesos argentinos ou bitcoins. Portanto, há uma

equivalência entre elas. Os valores de conversão, contudo, se alteram constantemente e não há nada fixo subjacente que essas diferentes moedas estejam medindo.

O que determina o preço de cada uma dessas moedas inclui uma diversidade de fatores, como oferta e demanda bens e serviços, relações diplomáticas entre os países, políticas governamentais de emissão de papel moeda, inflação, importação e exportação, confiabilidade dos governos, ação de especuladores e uma série de fatores históricos e contextuais difíceis de serem exaustivamente monitorados. Um bem produzido por um país pode ser internacionalmente valioso e, assim, tornar mais cara a moeda daquele país perante outras, mas esse valor pode cair abruptamente, por fatores incontrolláveis, levando também a uma queda naquela moeda.

Os bitcoins, por sua instabilidade e caráter descentralizado, podem trazer um exemplo interessante. São moedas e, portanto, conversíveis às outras moedas. Porém, são totalmente virtuais, não são controladas por nenhum governo, têm uma produção descentralizada em computadores espalhados pelo mundo. Alguns dizem que ela não é real por não possuir lastro. Recentemente, David Stockman, ex-diretor de orçamento da Casa Branca afirmou que os bitcoins não são dinheiro real, pois o dinheiro real tem que ser estável (UMPIERES, 2018). Essa afirmação, contudo, não parece correta. Afinal, se ele estivesse certo em relação ao motivo pelo qual os bitcoins não seriam dinheiro real, então uma grande parte das moedas estatais, inclusive o nosso real, também não seriam dinheiro real.

Porém, outros são mais otimistas, conferindo o lastro dos bitcoins a algoritmos matemáticos tão estáveis quanto o lastro físico em ouro (vale notar que as moedas nacionais contemporâneas não possuem lastro físico em ouro, seu lastro, entre fatores contextuais e políticos, tem a ver com o valor e a quantidade dos bens e serviços oferecidos por uma sociedade). Por exemplo, outro economista Ulrich (2014) foi bem confiante no valor dessas moedas. De acordo com ele, o que dá lastro a uma moeda é sua escassez relativa e as propriedades matemáticas dos bitcoins garantem essa escassez.

Qual o lastro do ouro? A escassez inerente de suas propriedades físico-químicas. Qual o lastro do papel-moeda fiduciário? A confiança de que governos não inflacionarão a moeda, apoiada em leis de curso forçado que obrigam os cidadãos a aceitar a moeda como pagamento. Qual o lastro do Bitcoin? Propriedades matemáticas que garantem uma oferta monetária, cujo aumento ocorre a um ritmo decrescente e um limite máximo e pré-sabido por todos os usuários da moeda. Após um bem ser empregado e reconhecido como moeda, seu lastro jaz na escassez relativa.

Observe-se, aqui, que se esse economista estiver certo, então a estabilidade do lastro nada tem a ver com a estabilidade do preço da moeda. Em outras palavras, se é verdade que as

propriedades matemáticas dos bitcoins lhe garantem um lastro estável, isso não foi suficiente para manter estável o seu preço, que já subiu 3.000% em dois anos, mas também já sofreu diversas quedas bruscas.

Essa digressão sobre lastro e conversibilidade das moedas e sobre bitcoins foi para servir como uma analogia para a indeterminação da tradução ou da interpretação, levando em conta que a interpretação envolve atribuição de estados mentais. As moedas, embora conversíveis umas às outras, possuem lastro instável que envolve uma série de fatores contextuais difíceis de serem controlados. Mesmo assim, por mais instáveis que sejam os reais que tenhamos no banco, podemos usá-los para comprar nossos alimentos. Aparentemente, é isso que lhes confere realidade, e não o seu lastro.

Em nossas vidas ordinárias, o dinheiro não precisa ser físico ou virtual, aparentemente, o importante é a possibilidade que ele carrega consigo para conversão em bens e serviços que necessitamos ou queremos consumir. O dinheiro de brinquedo que acompanha o jogo “Banco Imobiliário” não é dinheiro real porque não pode ser usado para irmos à feira comprar ovos nem pode ser investido em ações, nem convertido a nada que tenha esse tipo de utilidade. No mesmo sentido, o que confere realidade aos diferentes manuais de tradução é a capacidade que eles possuem para promover uma interação corrente e fluida.

Quanto às atribuições de estados mentais necessárias ao sucesso da interpretação, elas são reais na medida em que permitem explicar e prever adequadamente o comportamento alheio. Não é preciso haver algo fixo subjacente para que haja realidade frente à indeterminação. É preciso, sim que as versões concorrentes sejam relevantes para aquilo a que se propõem. Nesse quesito, quando se trata da atribuição de estados mentais, estamos diante de uma condição *sine qua non* para cada uma de nossas interações. Diante de duas atribuições de estados mentais divergentes e igualmente relevantes para a continuidade das interações, não é possível dizer de nenhuma delas que falte realidade ou que falte lastro.

Por essas razões, as diferentes moedas parecem trazer uma analogia mais apropriada para pensarmos nas atribuições de estados mentais do que outros sistemas de medição de coisas físicas, como polegadas e centímetros. A realidade da moeda, no fim das contas, é dada pelo seu poder de compra, ou seja, pela sua utilidade. Da mesma forma se dá a realidade das atribuições de estados mentais, elas são reais na medida em que permitem as diversas interações das pessoas. Outro ponto de analogia é que o valor da moeda se define contextualmente e envolve uma série de fatores instáveis das relações humanas.

É importante, porém, ressaltar um ponto de desanalogia para que se evitem confusões: As moedas são convenções culturais humanas. Já as atribuições de atitudes proposicionais,

embora estejam presentes em todas as nossas relações sociais e culturais, não se restringem a isso. Na realidade, antes que se possa formar qualquer relação social ou cultural, as atribuições de intencionalidade precisaram se fazer presentes nas situações de comunicação mais simples e pré-linguísticas. Por isso, não se deve compreender os autores interpretivistas como se atribuíssem um status meramente cultural aos estados mentais.

4.2.5. Dennett e a analogia com o dinheiro

Dennett (2005a) usa a analogia com o dinheiro para falar de algo um pouco diferente: ele quer fazer uma crítica aos filósofos que defendem a noção de *qualia* como propriedades intrínsecas, não relacionais, da experiência mental.

[C]ompare os *qualia* da experiência ao *valor* do dinheiro. Alguns americanos¹²² não conseguem tirar das suas cabeças que os dólares, diferente dos francos e marcos e ienes, têm *valor intrínseco* (“Quanto custa isso em dinheiro real?”). [...] Alguns participantes dos debates sobre a consciência simplesmente exigem, de forma direta, que as suas intuições sobre as propriedades fenomênicas sejam um ponto de partida inegociável para qualquer ciência da consciência (DENNETT, 2005a, p. 177-178, aspas e grifos do original, tradução nossa).

O que Dennett está denunciando aqui, é que quando se consideram os *qualia* como propriedades intrínsecas e não disposicionais, nada pode ser feito para esclarecer o seu significado. Entra-se no campo do mistério indecifrável. Além disso, ele pretende mostrar que o abandono de uma perspectiva ingênua e culturalmente autocentrada pode ajudar a trazer luz ao que parecia completamente misterioso.

O ponto-chave é a ilusão do “valor intrínseco” do dinheiro. Ele não existe, assim como não há um valor intrínseco por baixo das traduções rivais ou das diferentes interpretações. Provavelmente, também não haja valor intrínseco subjacente às diferentes atribuições de estados mentais. Os valores dos conjuntos das atribuições de estados mentais está na continuidade do sucesso das explicações e previsões comportamentais. Então, é preciso ter em mente que as atribuições de estados mentais não são feitas em forma de proposições isoladas, são conjuntos difusos.

Dennett (2017b) novamente trata da analogia com as moedas, e agora para pensar sobre outro assunto, embora relacionados. Agora, ele pretende pensar sobre a existência dos memes e estende essa analogia para a questão da existência da consciência e do livre-arbítrio. Ele quer defender a existência dos memes comparando-os com as cores. “Eu segui esse

¹²² Referindo-se especificamente aos estadunidenses, Dennett comete um deslize parecido com o que denuncia em seus compatriotas.

caminho por toda a minha carreira, insistindo que não apenas as cores são reais como também a consciência, livre arbítrio e dólares” (DENNETT, 2017b, p. 223, tradução nossa). Segundo ele, as pessoas tendem a não desconfiar da existência dos dólares, talvez por eles serem materializados em forma de cédulas e moedas. Os bitcoins, contudo, parecem mais ilusórios. Ocorre que são as instituições sociais que conferem realidade tanto aos dólares quanto aos bitcoins, eles são reais pelas mesmas razões. O engano não é dizer que os bitcoins ou os dólares sejam reais, mas atribuir a ontologia deles a algo que elas não são, isto é, moedas e cédulas, certificados de lastro em ouro ou qualquer outro tipo de objeto concreto.

Esse mesmo engano, para Dennett, deve ser evitado quando se trata da realidade dos estados mentais. Eles estão sujeitos à indeterminação, mas isso não significa que não sejam reais. Eles não são concretos, eles não podem ser localizados dentro nem fora do sujeito, mas dependem de uma série de relações complexas do sujeito com o mundo e com os demais sujeitos com quem ele interage. A analogia com as moedas pode evidenciar alguns pontos interessantes relativamente às atribuições de atitudes proposicionais: Ao contrário dos sistemas de medição, elas não medem nada fixo, o seu valor depende de uma série de fatores contextuais incontrolláveis. Por outro lado, já nascemos em sociedades cujas trocas são intermediadas pelo dinheiro. Como indivíduos, nada podemos fazer para alterar o seu valor. Ele está ligado a instituições complexas. O dinheiro é objetivo por ser altamente intersubjetivo. Sua objetividade não é como a das coisas físicas, é construída pelas relações de sociabilidade.

4.3. A questão ontológica: Realismo moderado

A ênfase dada por Dennett ao papel instrumental das atribuições de atitudes proposicionais não o torna um antirrealista acerca dos estados mentais. Ao contrário, em Dennett encontramos um realismo brando. Porém, o aspecto instrumental é anterior e dele deriva o realismo. A necessidade absolutamente vital de atribuição de estados mentais é, para Dennett, o que lhes confere realidade. É por meio da atribuição de estados mentais que somos capazes de nos comunicar e ela foi necessária para que nossa linhagem pudesse desenvolver a linguagem articulada, tendo uma base genética.

Para revisar, ao longo de vários milhares de anos, nós seres humanos viemos a apreciar os poderes das mentes individuais. Baseando-nos nos hábitos instintivos de todas as coisas vivas, nós distinguimos comida de veneno e, como outros organismos que se locomovem, nós somos extra sensíveis a animação (movimento guiado) em outras coisas móveis, e mais particularmente às crenças e desejos (informação e objetivos) que guiam esses movimentos, rastreando o melhor que

podemos *quem sabe o quê e quem quer o quê*, para guiar nossos próprios esforços de esconder e procurar. Esse viés nativo é a base genética para a postura intencional, nossa prática de tratar uns aos outros como agentes racionais guiados por crenças amplamente verdadeiras e desejos amplamente bem orientados. Nosso interesse ininterrupto nessas questões gerou a *psicologia de senso comum* na qual nós confiamos para dar sentido uns aos outros (DENNETT, 2017b, p. 380, grifos do original, tradução nossa).

Miguens¹²³, nessa mesma direção, também não aceitaria a denominação de “instrumentalista” à teoria dos sistemas intencionais de Dennett.

Ser *instrumentalista* acerca de alguma coisa é utilizá-la por motivos pragmáticos sem lhe conceder realidade ou importância fundamental. O instrumentalismo aqui em causa é um instrumentalismo quanto à linguagem mental. Mas a TSI¹²⁴ vem a mostrar não ser exactamente instrumentalista: o que ela faz é ligar o reconhecimento dos padrões da mentalidade ao ponto de vista de um intérprete, considerando que os padrões são reconhecíveis pelo intérprete devido à realidade do *design* que lhes subjaz. (MIGUENS, 2002, p. 22, grifos do original)

De fato, compreendido o instrumentalismo como a conjunção entre utilizar uma coisa por razões pragmáticas e não conceder a ela utilidade ou importância fundamental, a posição de Dennett acerca das atribuições de estados mentais está longe disso. Essas atribuições são absolutamente fundamentais, sem elas não seria possível explicar o mental. Porém, a razão pela qual atribuímos estados mentais aos outros e a nós mesmos, dentro do pensamento de Dennett, é pragmática de fato. As razões pragmáticas ganham uma força grande o suficiente para sustentar um realismo moderado, pois sem elas seria impossível seguir adiante. Nesse sentido, há algo de instrumentalista que sustenta o realismo na teoria de Dennett.

Certamente, o realismo moderado não vai conduzir a uma compreensão do mental como algo que existe além do mundo físico, nem como algo concreto ou fixo dentro do cérebro. Essa ideia se pode encontrar tanto em Dennett quanto em Davidson: o mental é uma maneira de descrever a realidade, pode haver outras maneiras, mas não serão apropriadas para determinados propósitos. Por isso, a realidade não poderia ser explicada completamente a partir de uma ciência da física. O vocabulário mental é imprescindível e irreduzível. Irreduzibilidade epistemológica não implica em irreduzibilidade ontológica, mas também não implica em antirrealismo.

Note-se que, dado que ambos aceitam a indeterminação aplicada à interpretação, ainda que a versão de Davidson da indeterminação seja mais branda (o que acaba por dar um pouco mais de peso ao seu realismo), em nenhum dos casos se pode levar a atribuição de estados mentais a um realismo forte. Quer dizer, não é possível encontrar um princípio de

¹²³ A leitura do trabalho de Sofia Miguens foi fundamental para o estabelecimento do ponto do realismo de Dennett e para que se pudesse perceber que não estamos diante de um autor instrumentalista.

¹²⁴ TSI: Teoria dos sistemas intencionais.

interpretação que seja capaz de definir qual é a atitude proposicional correta a ser atribuída em cada situação. Em primeiro lugar, porque as atitudes proposicionais não são atribuídas singularmente, mas em conjunto. Em segundo lugar, porque em tese¹²⁵, há uma série de conjuntos de atitudes proposicionais que poderiam ser atribuídos corretamente na interpretação do comportamento de alguém (ou de algo). Porém, alguns conjuntos seriam incorretos. Isso não é algo aleatório. Uma atribuição de atitudes proposicionais pode fracassar por não conseguir detectar um conjunto de crenças capaz de oferecer sustentação a determinados comportamentos.

Um cuidado que é sempre preciso ter em relação ao realismo de Davidson e de Dennett é o seguinte: ao se pensar em realismo das atitudes proposicionais, há uma tendência a considerá-las isoladamente. Porém, uma perspectiva atomista que coloca crenças individuais em confronto com a realidade para verificar se são verdadeiras é rechaçada por esses autores. As atitudes proposicionais fazem parte de um conjunto com limites difusos, elas só podem ser atribuídas sobre o pano de fundo de uma série de outras atitudes proposicionais.

4.3.1. Mais sobre a discussão entre Davidson e Dennett acerca do realismo

De que forma Davidson lida com a crítica de Dennett de que seu realismo está carregado demais? Por um lado, Dennett (1991a) pensa que, no processo de atribuição de atitudes proposicionais, a previsão do comportamento tem prioridade sobre a interpretação. A interpretação é importante por permitir a previsão. Davidson (1997/2001h), por outro lado, ataca o papel dado por Dennett à previsão comportamental. De acordo com Davidson, para que se façam previsões, é necessário haver uma teoria. Porém, no caso das pessoas, não seria possível construir teorias que permitam prever o que o outro vai fazer em seguida. Assim, Davidson não acha sequer que a previsão possa acontecer. Dennett, por sua vez, não compreende a previsão comportamental como algo proveniente de uma teoria. Para ele, a atribuição de estados mentais é uma habilidade não proposicional, mas baseada em um processo indutivo e natural de percepção de padrões. Percebe-se, então, que as posições dos dois autores têm mais divergências do que pareceria. Nesse ponto, eles parecem estar adotando pressupostos diferentes sobre as noções de teoria e de previsão. Porém, a principal

¹²⁵ Embora a indeterminação seja uma possibilidade teórica aberta em relação às atribuições de estados mentais, vale notar que as situações de interação que ocorrem na prática possuem uma quantidade muito grande de marcadores contextuais, de modo que o grau de indeterminação é bem pequeno nas situações de comunicação reais.

divergência entre eles não é essa. É, como Dennett (1991a) apontou, uma diferença de compreensão que ambos têm da indeterminação aplicada aos sistemas de atribuição de atitudes proposicionais.

Na concepção de Davidson, dois sistemas diferentes de atribuição de atitudes proposicionais podem estar de acordo com todas as evidências disponíveis, sem que haja uma forma de resolver qual deles seja o correto. Porém, se estão igualmente justificados a partir dessas evidências, eles são *compatíveis*. Dennett, por outro lado, leva a indeterminação a um nível mais profundo, sendo possível dois sistemas de atribuição de atitudes proposicionais *incompatíveis* estarem igualmente adequados às evidências.

Dennett confere uma realidade, ainda que abstrata, aos padrões de crença atribuídos durante a adoção da postura intencional. Porém, os padrões percebidos por pessoas diferentes podem ser outros e, nesse caso, não teria sentido buscar as atitudes proposicionais reais do sujeito. Davidson recusa essa ideia porque pensa que as atitudes proposicionais causam os comportamentos e os padrões abstratos de estados mentais, segundo ele, não poderiam desempenhar um papel causal¹²⁶.

Em relação à ontologia das atitudes proposicionais, Davidson não vê sua posição como sendo mais realista do que a de Dennett, pois os padrões de crença apresentados por Dennett possuem carga tanto ontológica quanto epistemológica. Para Davidson, a questão que pode ser colocada para se compreender a diferença de posição entre os dois interpretivistas não é qual dos dois defende uma teoria mais carregada em termos de realismo. A questão é, segundo Davidson “se há fundamentos objetivos para escolher entre hipóteses rivais” (DAVIDSON, 1997/2001h, p. 82-83, tradução nossa) diante de duas atribuições diferentes de atitudes proposicionais. Dennett diria que não há, mas Davidson pensa que pode haver, sim, pois a relação de triangulação entre sujeito, interlocutor e mundo pode estabelecer essa resposta. Segundo Davidson:

A fonte última (não o fundamento) da objetividade é, em minha opinião, a intersubjetividade. Se não estivéssemos em comunicação com os outros, não haveria nada em que basear a *ideia* de estar errado, ou conseqüentemente, de estar certo, tanto no que dizemos quanto no que pensamos. (DAVIDSON, 1997/2001h, p. 128, tradução nossa).

¹²⁶ Esta crítica de Davidson a Dennett focada nos possíveis poderes causais dos padrões de crenças se parece estranhamente com a acusação de epifenomenalismo feita por Kim (1993) a Davidson. E a resposta que pode aqui ser oferecida em defesa de Dennett se parece com a resposta dada por Davidson (1993/2005b) a Kim, isto é, falando de modo bem geral e resumido, que a forma como se descreve um evento não faz qualquer diferença para as relações causais nas quais esse evento está envolvido.

Em relação à comparação entre as atribuições de atitudes proposicionais e os sistemas de medição, é importante notar que há pontos de analogia e pontos de desanalogia. Um importante ponto de desanalogia, segundo Davidson, é que a medição comum usa números para captar os fatos relevantes. No caso das atitudes proposicionais, utilizam-se as frases. Assim, quando usamos um sistema de medição, entramos em acordo com nosso interlocutor acerca de qual sistema de medição estamos usando e com isso conseguimos estabelecer um ponto de partida extremamente simples. No caso das atribuições de atitudes proposicionais, as coisas não funcionam bem assim.

Você e eu não podemos chegar a concordar sobre a interpretação das nossas frases como um preliminar para usá-la para interpretar os outros, uma vez que o processo para se chegar a tal acordo envolve interpretação do mesmo tipo que pensávamos estar nos preparando para fazer. Não faz sentido perguntar por uma norma de interpretação comum, uma vez que a interpretação mútua provê a única norma que temos (DAVIDSON, 1997/2001h, p. 129, tradução nossa).

Na concepção de Davidson, os sujeitos entram em uma situação de comunicação munidos de uma teoria anterior acerca das intenções linguísticas uns dos outros. Porém, o ajuste comunicativo se dá por meio da teoria transitória criada durante a situação real e concreta de interação, onde os sujeitos em comunicação formam um corpo amplo de crenças compartilhadas¹²⁷. Certamente, isso traz para a interpretação das atitudes proposicionais um grau de indeterminação maior do que o encontrado nos sistemas de medição, como era de se esperar, dada a complexidade das relações humanas. Porém, o que se precisa notar no sistema de Davidson é que parece haver algo escondido sob as palavras, mas também revelado por elas e que se define na situação concreta de comunicação. Diferentes atribuições de atitudes proposicionais ao mesmo sujeito e sob as mesmas circunstâncias seriam apenas diferentes versões de um mesmo texto, sendo que esse texto não pode ter uma versão oficial. Esse aspecto relativamente estável do mental não existe no interpretivismo de Dennett.

Nesse sentido, o realismo de Davidson parece, sim, mais forte do que o de Dennett porque, para este, a realidade das atribuições de estados mentais se fundamenta na sua utilidade. Certamente, o termo “utilidade” tem um peso alto aqui. As atribuições de estados mentais são uma das coisas mais vitais para nós. Delas dependem todas as nossas relações de comunicação e toda previsão que fazemos dos comportamentos das pessoas. Davidson, por sua vez, fundamenta seu realismo nas relações de triangulação que faz com que, em algum ponto encoberto pela indeterminação, os significados possam se encontrar.

¹²⁷ Ver seção 2.3.

Há muito em comum entre os dois autores, mas enquanto Davidson dá ênfase à linguagem, Dennett enfatiza a evolução. Isso traz diferentes fundamentos ao peso que eles dão às atribuições de estados mentais. Para Dennett, o que caracteriza uma boa atribuição de estados mentais é o simples fato de ela obter sucesso na continuidade da comunicação. Por isso, duas atribuições contraditórias podem ser ambas corretas diante dos mesmos dados observáveis. Davidson não aceita isso. Para ele, diferentes atribuições de estados mentais para uma mesma situação só serão ambas corretas se forem versões diferentes de algo que vai pelo mesmo sentido.

4.3.2. *Uma compreensão mais recente do termo “ontologia” por Dennett*

A partir de uma leitura do livro mais recente de Dennett (2017b), é possível aprofundar essa concepção de ontologia que provém da utilidade epistemológica. Há duas noções importantes para isso: o conceito de “affordances” e a distinção feita por Sellars (1956/1997) entre imagem manifesta e imagem de ciência. Vejamos cada uma dessas noções para que seja possível a formação de um entendimento do que Dennett compreende por ontologia.

O termo “affordances” é um neologismo em língua inglesa, derivado de “afford”, um verbo que significa proporcionar, suprir ou dispor¹²⁸. Não há uma tradução precisa desse termo para a língua portuguesa. O termo foi usado primeiramente por Gibson: “As ‘affordances’ do ambiente são o que ele *oferece* ao animal, o que ele provê ou fornece, para o bem ou para o mal” (1979, p. 127, grifos do original, aspas internas nossas, tradução nossa). Mas o que o ambiente oferece a um animal são aspectos do mundo relevantes para aquele animal em algum sentido, conforme as possibilidades de relações entre aquele animal e o mundo circundante. Por exemplo, coisas que podem servir de alimento, coisas que podem aumentar ou diminuir seu sucesso reprodutivo ou coisas que ameaçam a vida e o bem estar.

Não se está tratando aqui de uma concepção subjetivista de fenomenologia. As “affordances” são partes dos padrões *reais* do mundo objetivo com os quais o organismo precisou aprender a lidar conforme suas possibilidades cognitivas. Porém, as possibilidades cognitivas de um organismo foram construídas a partir das relações que seus ancestrais

¹²⁸ Em notas sobre a noção de “affordance” feitas pelo prof. Paulo Abrantes (abr. 2018) para o grupo Mente, Linguagem e Evolução da UnB, algumas possibilidades de tradução são oferecidas: “o que pode ser aproveitado, explorado, apropriado, extraído (do meio ambiente); o que é concedido (pelo ambiente ao organismo) ou uma oportunidade (para o organismo)”. Essas possibilidades se revelam esclarecedoras para que possamos compreender o termo, eis a razão desta referência a material que não está publicado e foi elaborado apenas para uma discussão interna ao grupo.

tiveram com o meio. Em uma linguagem mais dennettiana, podemos dizer que as “affordances” são resultado de uma acumulação de projetos sem projetista. Elas são construídas ao longo da história evolutiva. Nesse sentido, se há um “como é ser” um organismo, esse “como é ser”, traz informações reais sobre a realidade externa, não está confinado no sujeito, ao contrário, é produto das relações de seus ancestrais evolutivos com o meio. O fato de ser real, não significa que seja um conhecimento perfeito. Nossa cognição é imperfeita por ser fruto de um processo local, contingente e imperfeito: o processo evolutivo.

Os organismos também alteram o ambiente em que vivem, construindo “affordances”. No caso humano, para Dennett, nossas “affordances” contém memes, o que faz com que nossas mentes sejam diferentes: “Graças a essa infestação de estruturas simbiotes de informação culturalmente evoluídas, nossos cérebros são capacitados para serem arquitetos inteligentes, de artefatos e de nossas próprias vidas” (DENNETT, 2017b, p. 370, tradução nossa).

Aqui entra a segunda ideia importante para compreendermos o que Dennett entende por ontologia: a distinção de Sellars entre *imagem manifesta* e *imagem de ciência*. Imagem manifesta é a imagem da realidade com a qual nós lidamos, as coisas que reconhecemos facilmente no mundo e com as quais estamos em um contato direto, as coisas às quais nosso discurso cotidiano costuma se referir. A imagem manifesta ancora as nossas relações de comunicação. A concepção de ontologia está ligada à imagem manifesta. “É esse o sentido no qual a “imagem” é manifesta: é óbvio para todos, e todos sabem que é óbvio para todos, e todos sabem *isso* também. Isso vem com a nossa língua nativa, é o mundo de acordo *conosco*” (DENNETT, 2017b, p.62, tradução nossa).

A imagem de ciência, por sua vez, é composta de coisas que não podemos ver, como moléculas, fótons e buracos negros. Precisamos aprender a imagem de ciência na escola, ela é muito diferente da imagem manifesta, aprendida em nossas relações sociais. Porém, conforme Dennett, a imagem de ciência tem como ancestral a imagem manifesta. Pela imagem manifesta, todas as pessoas podiam se entender. A partir dela se constituiu a imagem de ciência, a qual também compõe uma ontologia e também é capaz de ancorar certas relações das pessoas entre si (os cientistas) e com o mundo.

Essas duas versões do mundo são bastante distintas hoje, um pouco como duas espécies diferentes, mas elas já foram mescladas ou entrelaçadas em um único mundo ancestral do “que todos sabem”, isso incluiu toda a fauna local e flora, e armas e instrumentos e habitações e regras sociais, mas também duendes [“goblins”], e deuses e miasmas e feitiços que poderiam azarar sua vida ou garantir o seu sucesso caçando. Gradualmente nossos ancestrais aprenderam quais “coisas” derrubar de suas ontologias e quais novas categorias introduzir. Para fora foram as bruxas, sereias, gnomos e para dentro vieram os átomos, moléculas e germes. Os mais

antigos pensadores proto-científicos, tais como Aristóteles, Lucrecio e, muito mais tarde, Galileu, conduziram seus inquéritos sem fazerem uma distinção nítida entre a ontologia da vida quotidiana (a imagem manifesta) e a ontologia da ciência, mas eles foram proponentes corajosos de novos tipos de coisas, e as mais persuasivas delas foram popularizadas (DENNETT, 2017b, p. 62, tradução nossa).

Aqui, a concepção de ontologia de Dennett já pode ser delineada. O que ele entende por ontologia é o conjunto das coisas que estão disponíveis para o organismo conforme suas possibilidades cognitivas. No caso humano, é a nossa imagem manifesta que compõe a nossa ontologia, mas cada espécie animal possui uma ontologia específica. “Férias não estão na ontologia de um urso polar, mas neve está, e também estão as focas. Neve provavelmente não está na ontologia de um peixe-boi, mas hélices de motores de popa podem bem estar, junto com algas e peixes e outros peixes-boi” (DENNETT, 2017b, p. 60, tradução nossa).

Essas coisas são “affordances”, construídas por um processo evolutivo do qual o organismo simplesmente faz parte. Por isso, é preciso começar a tentar compreender esse processo e como ele deu origem às nossas “affordances”. No fim das contas, isso é uma maneira de se tentar compreender o ser humano como produto de uma relação direta com o mundo objetivo. A ontologia se compõe dos aspectos do mundo que estão disponíveis para que uma espécie mantenha relações apropriadas com seu entorno.

Dennett propõe uma inversão no ponto de vista a partir do qual os filósofos devem abordar a consciência: não mais a partir de dentro para fora, como se tem feito, mas sim de fora para dentro. A melhor maneira para que se possa compreender a consciência de fora para dentro, segundo ele, seria pela sua inserção dentro da história dos organismos vivos. É preciso compreender como ela passou a existir e qual seria a sua razão de ser. Uma noção que perpassa todo o livro é a de “raciocínios de livre flutuação” [“free floating rationales”], são razões para que as coisas existam, mas essas razões só fazem sentido quando estivermos tratando de organismos vivos. Elas não presumem uma inteligência para criá-las, por isso são de flutuação livre. “Por exemplo, há razões porque as colônias de cupins têm as características que têm, mas os cupins, diferentemente de Gaudí, não têm nem representam razões, e seus projetos [“designs”] excelentes não são produtos de um projetista inteligente” (DENNETT, 2017b, p. 52, tradução nossa). Só a consciência humana, na concepção de Dennett, é capaz de *representar* razões, mas as razões já existem livres na história da vida.

Tentar compreender a consciência humana de fora para dentro, na concepção de Dennett, é um hábito que promove uma ruptura entre os aspectos objetivo e subjetivo da realidade. Para Dennett, a imagem manifesta nos traz uma percepção do mundo exterior, não há uma realidade interior distinta da exterior. Segundo ele, nós experimentamos aspectos

estruturais do mundo que são importantes para nós, por exemplo, a doçura e a fofura. “O que é especial em relação a propriedades como doçura e fofura é que a percepção delas evoluiu para isso” (DENNETT, 2017b, p. 357, tradução nossa). Ou seja, não é a experiência privada que é especial, mas sim o fato de que essas propriedades são percebidas por nós com tanta intensidade porque essa percepção foi crucial para a sobrevivência e a reprodução de nossos ancestrais. Porém, ainda é preciso explicar por que nós temos a *impressão* de que há uma consciência composta de coisas inefáveis, intrínsecas (isto é, não disposicionais), privadas e diretamente apreensíveis para o sujeito e só para ele.

Na concepção de Dennett, as “affordances” dos seres humanos contém memes. Nós somos os seres capazes de representar as razões, elas também fazem parte de nossas “affordances” e precisamos lidar com elas o tempo inteiro, tanto para podermos compreender as outras pessoas quanto para justificarmos nossas ações. Essa necessidade nos levaria a uma apropriação de nossos pensamentos, criando a ilusão da consciência. Muitas vezes, precisamos dizer o que é que pensamos. Como é uma ilusão linguística, na concepção dele, é especificamente humana. Nós temos a necessidade de falar sobre nossas experiências. Nossa imagem manifesta faz com que olhemos para dentro a partir da adoção dos mesmos hábitos que costumamos adotar ao olhar para fora. Isso cria a ilusão de um mundo interno que é uma espécie de réplica do mundo externo feita de aparências. Criam-se, assim, os *qualia*, ilusões que nos permitem fazer referência ao mundo interior. Mas os *qualia* nada mais são, segundo essa concepção, que objetos intencionais. Dessa forma, Dennett compreende todos os estados mentais como sendo relacionais e, no caso dos *qualia*, eles são “affordances” que, entre outras “affordances”, permitem nossas relações com o mundo exterior.

4.3.3. Há um tipo de dualismo nesses autores?

Já foi mencionada uma crítica de Kim (1993) a Davidson segundo a qual este não conseguiria evitar o epifenomenalismo. Também se poderia entender essa crítica como uma acusação de dualismo ao pensamento de Davidson. A crítica provém da impossibilidade, preconizada por Davidson, de redução do mental às explicações causais envolvendo leis estritas na realidade física.

Miguens lança uma crítica a Dennett que parece possuir algum parentesco com crítica de Kim a Davidson. Segundo ela:

[A] TSI¹²⁹ pretende ser ao mesmo tempo uma teoria fiscalista e uma teoria quineana da interpretação e a verdade é que um certo dualismo persiste nas teorias quineanas da interpretação: tanto quanto estas propõem que o mental é *apenas uma interpretação* de algo que é *realmente físico*, elas retêm por princípio uma distinção entre o mental e o resto. É certo que não se trata de um dualismo imaterialista ou solipsista mas trata-se ainda assim de um dualismo, de resto perfeitamente notório nas discussões acerca de sintaxe e semântica, razões e causas, vocabulário intencional e vocabulário físico que povoam a filosofia da mente (MIGUENS, 2002, p. 481, grifos do original).

Essa crítica poderia perfeitamente se aplicar a Davidson, pois ele também é adepto de uma teoria quineana da interpretação. Há uma diferença importante entre a crítica de Miguens e a de Kim. Miguens reconhece que a acusação de dualismo não é profunda, ela não desce ao nível da ontologia, mas é um dualismo acerca de modos de explicação da realidade. De fato, Davidson entende que não é possível dar sentido aos eventos mentais por meio de uma explicação restrita à física como disciplina. Dennett também pensa isso. Então, embora haja apenas um nível ontológico, há mais do que um nível epistemológico.

Se usamos o termo “dualismo” para as questões epistemológicas, Dennett e Davidson serão considerados dualistas? Em uma concepção muito suave do termo, eles são dualistas por considerarem as mentes humanas como sendo um tipo muito específico de mente, diferente das mentes dos outros animais. Porém, há um outro sentido em que talvez Davidson possa ser considerado dualista e Dennett, não. Para Davidson, há dois tipos de vocabulários para se descrever a mesma realidade: o vocabulário físico e o mental. Mas Dennett incorpora diversos níveis de explicação para a realidade. Ao longo do tempo, a proposta interpretivista de Dennett teve tempo para se sofisticar. Muitas coisas aconteceram: importantes desenvolvimentos em inteligência artificial e nas ciências cognitivas, além de mudanças na forma como o pensamento filosófico passou a se relacionar com o desenvolvimento científico.

Dennett agrega os desenvolvimentos das mais diversas disciplinas científicas em seus estudos acerca do mental. O resultado é que, para Dennett, um dualismo epistemológico seria algo simples demais. Embora a realidade seja uma só, nosso olhar sobre ela é estratificado. Em outras palavras: não há apenas duas formas de explicá-la, mas diversas. Em primeiro lugar, porque há diversas disciplinas científicas cujos vocabulários são irreduzíveis. Isso já cria uma série de camadas explicativas. Há as explicações da física e há as da química. Também há as explicações da biologia, essas em vários níveis: nível celular, nível dos tecidos, órgãos, desenvolvimento ontogenético e filogenético, perspectivas evolucionistas mais amplas, etologia comparada e assim sucessivamente. Aspectos sociais e históricos, sobretudo, devem ser levados em conta para que se possa tentar explicar a mente humana.

¹²⁹ TSI: Teoria dos sistemas intencionais

A compreensão do sujeito, na concepção de Dennett, passa pelo olhar que o sujeito tem sobre si, pelas histórias que os outros contam acerca dele e também pelo que as diversas ciências têm a dizer. Quanto mais de perto se olha para o sujeito, conforme pensa o autor, mais ele é fragmentado. Quanto mais afastado estivermos dele, mais o sujeito parecerá integrado. Isso significa que o olhar do sujeito sobre si mesmo é distante, afastado. Ele tem a ilusão de ser um sujeito integrado, uma espécie de fantasma encapsulado em seus limites de emoção, razão, experiências e memórias. Porém, tudo isso está em um fluxo constante e difuso, subjacente à autopercepção que o sujeito tem de si.

Em seu livro mais recente, Dennett (2017b) procura diluir a distinção entre razões e causas, argumentando que as razões podem ter surgido a partir das causas. Há dois tipos de “porquês” que ele procura diferenciar: o “para quê” e o “como”. Então, as explicações para os seres pré-biológicos e para a origem dos primeiros replicadores seria em termos de “como”, como eles vieram a surgir. Para os replicadores, então, talvez se possa gradualmente começar a tentar entender o “para quê”. Para que, por exemplo, determinada estrutura está em determinado organismo? “Assim como não há o Prima Mamífero - o primeiro mamífero que não tinha um mamífero como mãe - não há Prima-Razão, a primeira característica da biosfera que ajudou algo a existir porque o fez melhor em existir do que a ‘concorrência’” (DENNETT, 2017b, p. 49, aspas duplas do original substituídas por aspas simples). A transição das causas às razões é gradual e não encontra um momento histórico específico para acontecer. Na concepção do autor, não faz sentido dizer que as causas fazem parte do domínio da física e as razões, do domínio do mental. Razões e causas fazem parte da compreensão da realidade como um todo, mas antes de haver replicadores não havia razões e não houve uma razão para que eles surgissem.

Dennett adota uma postura adaptacionista. O adaptacionismo olha para as estruturas de um organismo buscando compreendê-las como adaptações, com o cuidado de buscar verificar também as que não são adaptações¹³⁰. Nesse sentido, ele considera esclarecedor olhar para as estruturas buscando suas *funções*. Assim como há razões para nossos comportamentos, o autor defende que há razões também para muitas outras coisas que não possuem razões nelas mesmas, ou que não buscam razões. Essas razões estão na mãe natureza. A pergunta pela razão, no fim das contas, não se restringe aos humanos. A natureza tem razões para agir, mesmo não tendo noção disso. “Darwin não extinguiu a teleologia, ele a naturalizou” (DENNETT, 2017b, p. 51, tradução nossa).

¹³⁰ É preciso haver uma cautela em relação ao adaptacionismo. Esse debate foi levantado por Gould e Lewontin (1979), mostrando que seria “panglossiano” olhar para todas as estruturas como adaptações.

Mesmo assim, Miguens está correta ao atribuir a Dennett a ideia de que o mental é algo completamente diferente de todo o resto no que diz respeito às mentes humanas. Embora as razões sejam amplamente disseminadas, nós somos a única espécie capaz de se engajar em um *controle* das razões. Nós perguntamos as razões que as pessoas tiveram para agir e avaliamos essas razões. O que nós temos de diferente é a capacidade de apreciar as razões. Para Dennett, essa capacidade é elemento fundante da vida social. Se alguém não tiver essa capacidade, não será considerado portador de responsabilidade moral e não poderá ser submetido às mesmas normas sociais da maioria.

Apesar de Dennett apresentar a mente humana como algo muito diferente das outras mentes, seus argumentos são sempre fundados na evolução e no algoritmo evolutivo. Há diversos tipos de mentes¹³¹, cada uma terá características específicas. Mas o que Dennett pretende dissolver é a compreensão do mental como algo que permanece indefinidamente como um mistério. Ainda que seja diferente, isso não significa que não possa ser compreendida. Há um gradualismo na explicação de Dennett. Assim, a Dennett seria atribuída a ideia de estratificação, não apenas de dualismo epistemológico. Isso não desce ao nível da ontologia. O mesmo em relação à acusação de dualismo para Davidson. Ele realmente é irreducionista em relação ao mental. O vocabulário mental, para ele, realmente não se reduz ao físico. Mas isso não é uma acusação, é o que Davidson quer defender, mesmo. Isso não é epifenomenalismo, o dualismo não desce para a ontologia.

Apesar dessas explicações, é preciso lembrar que, para ambos os autores, o ser humano tem algo que o faz muito diferente dos outros animais. Está claro, em Davidson, que é a linguagem articulada, algo que constrói uma mente especificamente humana. Para Dennett, nossas mentes são diferentes porque nós temos “ferramentas para pensar”, sendo as palavras faladas as mais fundamentais. Mesmo assim, Dennett apresenta uma explicação para o surgimento das mentes humanas, culturais e linguísticas, que é gradualista, evita grandes lacunas.

4.4. A questão epistemológica: Falsas crenças não andam sós

Foi dito que, para Davidson, as nossas crenças são majoritariamente verdadeiras, além de compartilharmos de um fundo massivo de crenças compartilhadas. O fundamento para isso está na triangulação, o processo pelo qual nós nos tornamos capazes de comunicação linguística. Então, o que dá sustentação a esse direcionamento da maior parte de nossas

¹³¹ Ver: seção 3.1.7.

crenças para a verdade, em Davidson, é a linguagem. Seus argumentos se aplicam, de modo geral, apenas às atitudes proposicionais. No caso de Dennett, há uma tendência a que os organismos reajam ao mundo mediante um sistema cognitivo construído de modo a permitir relações apropriadas entre esse sistema e o mundo, por mais imperfeito que seja esse sistema cognitivo, ele dá informações relevantes sobre o mundo, conforme as “affordances” daquele organismo.

Aqui, é importante deixar claro que a própria noção de crença para Dennett tem uma diferença chave em relação a Davidson. É que a crença, para Dennett, não é necessariamente linguística. Em primeiro lugar, porque não há bases para fundamentar a afirmação de que um cão, por exemplo, não possui crenças. Certamente, há crenças, como por exemplo, crenças sobre a democracia, que dependem da posse de uma linguagem simbólica complexa. Mas, na concepção de Dennett, não há nenhuma razão definitiva para que possamos dizer que o cão não tem a crença, por exemplo, de que seu dono pode lhe dar ração. Em segundo lugar, porque nós, os que podemos falar sobre algumas de nossas crenças, parecemos ter muitas crenças as quais não somos capazes de apresentar verbalmente e das quais sequer estamos cientes (DENNETT, 1976, p. 181). Por outro lado, para Dennett, a implementação de uma mente consciente depende de uma linguagem articulada. Nesse sentido, embora compartilhemos com outros animais o fato de termos crenças, a mente humana é muito diferente das outras mentes. “Nós sabemos que existem bactérias, cachorros não, golfinhos não. Mesmo as bactérias não sabem que há bactérias.” (2017b, p. 3, tradução nossa). A diferença é que a mente humana possui a ferramenta da linguagem.

Ele toma como pressuposto que as crenças com valor de sobrevivência são as que melhor aproximam o indivíduo do real, pois as crenças são ferramentas para nossas interações com o mundo. “A psicologia de senso comum simples nos diz que, uma vez que as pessoas usam suas crenças para selecionar e guiar suas ações, as crenças verdadeiras são sempre melhores que as crenças falsas” (McKAY; DENNETT, 2009, p. 508, tradução nossa). Essas ferramentas foram desenvolvidas ao longo de um processo que projetou os sistemas para se voltarem para a verdade de modo imperfeito, gerando crenças rudimentarmente justificadas. “Crenças embasadas [“grounded beliefs”] são simplesmente crenças que estão apropriadamente fundadas na evidência e nas crenças existentes” (McKAY; DENNETT, 2009, p. 494, tradução nossa).

Porém, como se pode explicar o fato de que carregamos crenças contraditórias entre si? Essas contradições são permitidas até certo ponto em um sistema intencional. De acordo com Miguens, em um sistema intencional, (2002, p. 167), “alguma inconsistência é sempre

permitida e apenas a ilogicidade em grande escala indica ou um defeito tão sério no sistema que o desqualifica como crente ou um defeito nas hipóteses de interpretação”. Quer dizer, como a caracterização das atitudes proposicionais de um sistema intencional é feita de maneira holista, ao encontrar um grau muito elevado de contradição, o intérprete deve rever sua prática, algo está errado. Essa é uma ideia que provém de Quine e está presente em Dennett.

Dennett (1978/2006a) diferencia opinião e crença para poder explicar a acrasia¹³² e o auto-engano. Segundo ele, as crenças seriam causas diretas do comportamento, juntamente com os desejos. Por isso, poderiam ser atribuídas também aos animais que não possuem linguagem, sendo detectadas a partir de sua função na decisão comportamental. Porém, as opiniões dependeriam da linguagem. Assim, a acrasia e o auto-engano são evidenciados pelo comportamento contrário ao juízo, algo que, para Dennett, os animais não humanos não poderiam fazer. Outro aspecto importante na compreensão dessas falhas humanas seria a necessidade que temos de mudar de opinião diante de argumentos. Podemos ter o nosso comportamento influenciado por crenças inculcadas em nós, mas que nunca paramos para avaliar. Por isso, a importância da mudança de opinião na constituição de um agente de responsabilidade moral.

Um problema com que é preciso lidar aqui é entender o fato de que nós, humanos, somos capazes de ir fundo em elucubrações, paranoias, histórias infundadas, mentiras e “fake news”. Algo que deve ser levado em consideração para a compreensão, dentro do pensamento de Dennett, de como as falsas crenças encontram território fértil nas mentes linguísticas, é sua proposta de que há um processo de evolução cultural em cena, paralelo ao processo de evolução biológica. Esse processo envolve um ambiente diferente e uma velocidade diferente, além de diferentes entidades replicadoras, mas, em um nível altamente abstrato, é regido pelo mesmo algoritmo evolutivo. Os dois processos se relacionam, mas uma vez que as entidades replicantes são diferentes, eles podem, em alguns momentos, entrar em antagonismo. Tal antagonismo talvez possa favorecer os memes, no sentido de lhes proporcionar um maior sucesso reprodutivo. Em outras palavras, talvez algumas crenças falsas se proliferem não por aumentar o sucesso reprodutivo dos animais linguísticos que carregam essas crenças, mas sim, por terem encontrado, elas mesmas, uma alta taxa de sucesso reprodutivo no ambiente cultural.

Eis a importância de se estabelecer uma diferença entre crenças não linguísticas e linguísticas: as não linguísticas direcionaram os comportamentos de nossos ancestrais,

¹³² A incapacidade de agir de acordo com seu melhor juízo.

abrindo passagem para o surgimento de organismos suficientemente complexos para desenvolverem a linguagem articulada. Mas isso também abriu caminho para que um novo processo evolutivo entrasse em cena e direcionasse o comportamento humano, não apenas para a replicação dos genes como também para a replicação dos memes. Essa é a posição de Dennett que, embora discutível¹³³, tem a vantagem de trazer uma luz para o fato de que os processos culturais muitas vezes parecem nos colocar contra nossos próprios interesses de um ponto de vista biológico.

No fim das contas, a evolução não está preocupada com a verdade, mas sim com o sucesso reprodutivo, essa é uma lição que fica bem evidente no pensamento de Dawkins (1976). Por exemplo, segundo McKay e Dennett (2009), há pesquisas recentes mostrando que muitas pessoas tem uma tendência a se superestimarem, considerando-se mais inteligentes e competentes do que seus colegas. Porém, ao contrário do que se poderia imaginar, essa falsa crença acaba por se revelar benéfica para aqueles indivíduos que a têm, dando-lhes autoconfiança e permitindo que alcancem melhores postos de trabalho. Nesse sentido, os autores conjecturam que a evolução memética possa ter favorecido algumas crenças falsas.

Mas isso explica apenas parcialmente a questão, pois pegamos de Davidson a lição de que, ao aprender uma linguagem, partimos das relações de triangulação com o mundo e o interlocutor. Então, por mais estranhas que pareçam algumas das fantasias humanas, elas se constroem sobre um tripé bem apoiado no chão. Como se dá a passagem dessas crenças formadas por um processo tão confiável para uma grande quantidade de crenças absurdas que, frequentemente, lhe sobrevém?

Qual o sentido, por exemplo, de a maior parte dos grupos humanos compartilharem de mitos religiosos que são completamente diferentes de uma comunidade para outra? Provavelmente, a religião traga um conforto às pessoas em momentos difíceis da vida, mas também tem “justificado” uma série de atos fundados na ignorância e prejudiciais aos fiéis. Há muitas pessoas que direcionam amplamente suas vidas a partir de preceitos religiosos. Mas a própria religião é um produto da cultura, alicerçada nas relações de triangulação, tão firmes e bem plantadas. Como isso é possível¹³⁴?

¹³³ Ver: seção 3.4.3.

¹³⁴ Dennett é um militante ateu. Segundo ele (2006b), a religião exige moralmente das pessoas apenas boas intenções no coração, levando-as a uma maior auto-indulgência perante suas próprias falhas morais. Parece uma ideia correta, mas não explica como essa tendência tão forte ao engano poderia se estabelecer evolutivamente. Essa questão será deixada em aberto neste trabalho. A resposta, de acordo com Dennett, deve ser darwinista, considerando o quanto a religião é amplamente disseminada entre os humanos. Contudo, essa resposta darwinista não precisa necessariamente se basear em aspectos biológicos, já que a própria cultura, para o autor, é um processo darwinista (DENNETT, 2003a). Vale notar, contudo, que uma explicação darwinista para a religião ainda seria parcial, pois não nos permitiria compreender as crenças falsas de modo geral.

Permanece a dificuldade de compreender como pode um sistema intencional qualquer, seja linguístico ou não linguístico, ter crenças falsas se o valor de sobrevivência das crenças aparentemente as direciona para a verdade. A compreensão que Dennett possui de crença não permanece estável em sua obra. “Uma crença é um estado funcional de um organismo que implementa ou incorpora o endosso daquele organismo a um particular estado de coisas como real” (McKAY; DENNETT, 2009, p. 493, tradução nossa). Uma crença falsa, então, seria aquela que endossa um estado de coisas o qual o organismo não consegue alcançar.

Porém, é preciso levar em consideração que as crenças são habilidades cognitivas que surgem na história dos seres vivos posteriormente a outras características adaptativas. Em geral, as características adaptativas se desenvolvem a partir da relação do sistema com situações ou problemas localizados no tempo e no espaço. Um aspecto adaptativo em um ambiente pode perder esse valor em outro ambiente sem necessariamente levar aquela espécie à extinção, uma vez que ela pode ser protegida por outras características que se adaptam bem ao novo ambiente, ou pode ter sofrido mutações em outras características que lhe permitiram a reprodução nessas novas condições. “A evolução [...] não é um processo de projetos perfeitos, mas é sujeita a limitações econômicas, históricas e topográficas” (McKAY; DENNETT, 2006, p. 509, tradução nossa).

Diversas situações contingentes e locais, portanto, podem explicar, antes mesmo do surgimento das crenças falsas, qualquer característica cognitiva de um organismo que pareça equivocada. A evolução, afinal, é um relojoeiro sem propósito. É preciso ver a função que uma característica adaptativa desempenhava historicamente, pois isso pode ter se perdido ao longo das mudanças ambientais. De todo modo, tomando-se o conjunto das características cognitivas de uma espécie, provavelmente haverá uma tendência a que elas levem aquele animal a uma relação apropriada com o ambiente, ou seja, uma relação que aumente suas chances de sobrevivência, pois se não for assim, os indivíduos passarão a morrer antes de deixar descendentes. Novamente, isso não impede que algumas das características cognitivas daquele animal o levem a enganos e erros de funcionamento. “O fato de que nós não estejamos presentemente equipados com sistemas de formação de crenças à prova de falhas não conta contra uma perspectiva evolucionária. Isso porque a evolução não produz necessariamente sistemas otimamente projetados” (McKAY; DENNETT, 2009, p. 497, tradução nossa). As imperfeições do processo evolutivo, assim, produzem naturalmente erros na representação da realidade.

Isso pode ser transportado para as crenças. Nós, os animais linguísticos, produzimos muitas crenças falsas, histórias absurdas, acreditamos em mentiras perpetuadas ao longo de

gerações, somos capazes até de fechar os olhos para não vermos a verdade. Há as falsas crenças criadas pela mente, há os erros do funcionamento do cérebro, há as imperfeições do sistema cognitivo produzido por um processo imperfeito (a evolução não é onisciente, onipresente nem onipotente). Por isso, pensa Dennett: “Uma boa teoria da consciência *deve* fazer uma mente consciente parecer uma fábrica abandonada, cheia de máquinas sussurrando e ninguém em casa para supervisionar, ou aproveitar, ou testemunhar” (DENNETT, 2003a, p. 16, tradução nossa). Porém, a fricção de nossas atitudes proposicionais com as das outras pessoas e com o mundo vai ajustando as coisas de modo que a quantidade de crenças verdadeiras seja maior, garantindo a possibilidade de comunicação.

Diante de todos os erros, contingências e imperfeições, pode-se ainda argumentar que há dois processos assegurando que nosso comportamento seja guiado por crenças majoritariamente verdadeiras, ambos processos imperfeitos. Podemos chamar ao primeiro processo de dennettiano e ao segundo de davidsoniano. O primeiro é o processo evolutivo. Ele se direciona não à verdade, mas à reprodução. Porém, se produzisse sistematicamente estados cognitivos massivamente enganadores, conduziria os organismos à morte, não à reprodução. Imagine, por exemplo, um organismo que sistematicamente acredite ser alimento o que é veneno. Nesse sentido, a evolução leva o organismo a um direcionamento bastante imperfeito e instável aos ajustes cognitivos frente ao mundo. Nessa perspectiva, a verdade não é totalmente objetiva nem totalmente subjetiva, ela é resultados dos ajustes históricos entre organismos e mundo ao longo da história da vida.

O segundo processo funciona apenas para os sistemas linguísticos. O processo de aprendizagem da linguagem e a maneira como as crenças linguísticas se apoiam mutuamente faz com que, tomadas de forma global, as crenças do sujeito tenham uma tendência a serem majoritariamente verdadeiras e a falsidade, no âmbito da linguagem, só é possível sobre esse pano de fundo. Então, por mais que nos envolvamos com os mais diversos tipos de elucubrações e mentiras, há uma base sobre a qual elas se erguem e essa base está enraizada na realidade.

Nesse sentido, vale lembrar que não é possível individuar crenças para verificar se elas estão ou não em um sujeito. Ao lhe atribuir uma crença, uma série de outras crenças são atribuídas concomitantemente, dando suporte à interpretação em um sentido amplo, e não às crenças individuais. Por exemplo, Maria crê, com base em uma série de experiências passadas, que sua amiga Carla é honesta. Porém, Carla guardou na própria bolsa a carteira de Maria sem que esta percebesse. Ao descobrir o ocorrido, Maria atribui a Carla a crença de que guardou a carteira por achar que pertencia a si (Carla) e não a Maria, considerando que as carteiras de

ambas era da mesma cor. Ao comentar o caso com sua mãe, uma nova hipótese é lançada: Carla pode ter guardado por pura distração, estava absorvida por algum pensamento e guardou a carteira de Maria por engano, sem refletir sobre seu comportamento. Nesse caso, Dennett não considera possível descobrir o que tenha de fato ocorrido na mente de Carla. A razão para isso é que não é possível isolar uma crença para avaliá-la.

O holismo do mental, então, faz com que nossas crenças precisem se acomodar dentro de um sistema de crenças. O modelo dos rascunhos múltiplos, do qual falamos no primeiro capítulo¹³⁵, permite que nossa mente “edite” as crenças que parecem destoantes, acomodando-as de alguma forma ao sistema global de crenças. Na maior parte das vezes, os relatos que os sujeitos fazem de seus estados internos correspondem a crenças provocadas pelo funcionamento normal de seu sistema cognitivo. Porém, há circunstâncias, como o caso do *déjà vu*, em que a acomodação se torna imperfeita. Nesse caso, segundo Dennett (2001), estaríamos diante de um subproduto da atividade cerebral que gera a ilusão do *déjà vu*. Dennett recorre a pesquisas já realizadas acerca desse fenômeno. Um erro do sistema cerebral subjacente seria responsável pela sensação de *déjà vu*. A impressão de que algo “mágico” está acontecendo decorre do fato de que o sujeito não tem acesso aos processos físicos mais simples que subjazem à sua percepção. Mas no fim das contas, não é uma “mágica real”, e sim um truque de mágica (DENNETT, 2003b). Assim como a evolução, nossa mente é imperfeita e produz erros, mas funciona incrivelmente bem.

4.4.1. Truques cerebrais

Essa impressão de mágica do *déjà vu* é produzida porque a experiência rompe bruscamente com nossas expectativas. Porém, na concepção da Dennett, a própria experiência consciente, no fim das contas, é um truque do cérebro, mas é um truque com o qual estamos acostumados e por isso não parece mágico. O método para compreendê-la, por isso, seria o mesmo método utilizado para se compreender um truque de mágica: verificar os artifícios simples que estão por traz daquela ilusão. A mágica real é a ilusão do expectador, nada que não possa ser compreendido, desde que se desvende o truque.

O truque de mágica produz a impressão de uma consciência única e integrada, a impressão de experiências puramente subjetivas e infáveis e de atitudes proposicionais isoladas. Essa impressão é criada nas mentes linguísticas, devido à sua complexidade. Ela cria uma ilusão integradora dos processos mentais, quando eles são, na realidade, altamente

¹³⁵ Ver seção 1.1.1.

difusos, estão em um fluxo constante e incontrolável, mudam o tempo inteiro em uma velocidade muito pequena para nossas experiências macroscópicas. O tempo e o espaço dos processos mentais são microscópicos, mas a natureza nos dotou com um sistema cognitivo necessário para controlar o tempo e o espaço em uma escala macroscópica. Assim alcançamos um excelente grau de previsibilidade e explicação em relação aos eventos exteriores.

Dennett mostra que uma crença falsa não pode ser inserida isoladamente dentro do sistema de crenças de uma pessoa. É um fato que as memórias das pessoas sofrem alterações frequentes e há muitos casos em que memórias falsas são inseridas (LOFTUS, 1997), provocando uma série de consequências comportamentais naquele indivíduo (BERKOWITZ et al, 2008). Mas essa inserção consiste em um processo no qual uma série de detalhes são construídos e a falsa memória acaba se acomodando junto com outras memórias de modo razoavelmente harmônico, em um processo que, para Dennett, se pareceria com uma enxurrada de revisões editoriais. Uma memória falsa inserida em um sistema não é uma atitude proposicional isolada. Por mais estranha que ela seja, precisará se acomodar com muitas outras crenças daquele indivíduo.

4.5. Questões de tendência ontológica e questões de tendência epistemológica

Na introdução desta tese foram propostos dois grupos de questões: questões de tendência ontológica e questões de tendência epistemológica. Neste momento, estamos em condições de retornar às questões. Após essa busca de uma compreensão ampliada do interpretivismo de Davidson e de Dennett, relacionando essa posição a outros aspectos do pensamento de cada um desses autores na tentativa de esclarecê-las, algumas das questões inicialmente colocadas parecem perder um pouco o sentido por partirem de uma concepção fragmentada das realidades objetiva e subjetiva. Vejamos, então, cada uma das questões apresentadas.

A primeira questão do grupo de perguntas de tendência ontológica era essa: Numa perspectiva interpretivista, os estados mentais têm uma existência no sujeito ou são apenas algo que se atribui de uma perspectiva externa, sem que haja nada do lado de dentro? O problema dessa pergunta é apresentar uma disjunção excludente. Apenas depois de desenvolvido todo esse estudo acerca desses autores é que se pode perceber que os estados mentais de um sujeito podem ser capturados a partir de uma perspectiva externa porque foram construídos a partir das relações desse sujeito com o mundo. Assim, embora sejam estados do sujeito, não estão dentro do sujeito, assim como não estão fora do sujeito porque não possuem

uma localização espacial, sua existência não é concreta. Eles são projetos e construções das relações entre sujeito e mundo.

A partir dessa pergunta, surge a pergunta sobre se os estados mentais são reais conforme a perspectiva desses autores. São reais, sim, na medida em que fundamentam todas as nossas relações interpessoais e muitas das nossas relações com o mundo objetivo. A realidade deles provém de sua força epistêmica, de sua capacidade de oferecer um caminho filosófico alternativo onde o ceticismo acerca das outras mentes e do mundo exterior não se apresentam. A proposta de se estudar o mental não a partir do ponto de vista do sujeito, mas sim pela perspectiva das relações entre sujeito e realidade objetiva e intersubjetiva traz uma realidade objetiva para os estados mentais, ainda que não seja uma realidade concreta.

Essa pergunta nos remete à crítica que John Heil (1998) faz ao interpretivismo. Para Heil, a comparação dos estados mentais a sistemas de medição feita por uma perspectiva interpretivista os levaria a um status de meras convenções culturais. Ao contrário da leitura de Heil, o que Davidson pretende ao comparar as atribuições de atitudes proposicionais a sistemas de medição é enfatizar que há algo fixo subjacente aos diferentes sistemas de medição, isso que é real nas atribuições de estados mentais. Dennett, por sua vez, embora não defenda a existência de algo fixo, ancora a realidade dos estados mentais, assim como das atribuições, no processo evolutivo, trazendo-lhes um peso muito maior do que meras convenções culturais. A postura intencional, inclusive, possui uma base genética. Na concepção de Dennett, ela é inata.

Portanto, contrariamente à leitura de Heil, nenhum dos autores interpretivistas (nem Davidson, nem Dennett) dá às atribuições de estados mentais o status de meras convenções culturais. Ao contrário, sem elas, não haveria sequer nenhuma convenção cultural, pois atribuições de estados mentais são anteriores a qualquer situação de comunicação linguística. O ponto da analogia com os sistemas de medição proposta por Davidson é outro, mas Heil coloca o olhar justamente sobre o ponto de desanalogia. A analogia é para mostrar que, embora haja diferentes sistemas para medir alguma coisa, há algo subjacente sendo medido. A desanalogia é que os sistemas de medição são convenções culturais, mas as atribuições de atitudes proposicionais, não.

Vale notar que a estratégia de derivar o realismo moderado da utilidade já havia sido utilizada por Quine (1960/2010), mas para tratar de números e classes. Na discussão sobre nominalismo e realismo acerca das classes, Quine acaba por adotar uma posição realista fundada sobre a eficácia desses conceitos para as ciências. De acordo com ele, os números e as classes de objetos devem também ser considerados objetos, justamente por sua eficácia.

Conforme Smart (1969, p.5, tradução nossa): “Números e conjuntos são postulados da mesma maneira como neutrinos e elétrons. Em ambos os casos nossos postulados formam parte da base para uma teoria bem testada.”

No mesmo sentido, Putnam (1975) oferece o argumento do milagre em favor de uma abordagem realista da ciência. De acordo com esse argumento, a melhor explicação para o sucesso das teorias científicas na explicação e previsão de uma série de fenômenos é que elas são verdadeiras. Caso contrário, seria um milagre, uma coincidência extremamente improvável que as teorias científicas produzissem o sucesso que produzem.

Observe-se que a estratégia de Dennett para sustentar o realismo das atribuições de estados mentais é a mesma que Quine usou para tratar das classes e números e muito parecida com a estratégia usada por Putnam para tratar das teorias científicas. Porém, Dennett não precisa aceitar o realismo das teorias científicas nem das classes e números para aceitar o realismo da psicologia de senso comum. Isso porque a psicologia de senso comum é mais fundamental do que as teorias científicas e do que as classes e números. A psicologia de senso comum está na base de todas as nossas interações sociais e já estava presente no ancestral comum entre nós e os grandes símios. Sem ela, não haveria linguagem, muito menos números, classes ou teorias científicas.

A terceira pergunta era sobre se uma postura interpretivista seria capaz de reconhecer estados mentais pré-linguísticos, em especial, nos animais não humanos. Ora, Davidson reconhece que há uma triangulação pré-linguística. Dennett reconhece que os animais não humanos também possuem sistemas cognitivos que propiciam o reconhecimento de padrões reais no mundo. Alguns desses animais são capazes, inclusive, de adotar a postura intencional. Porém, ambos os autores consideram a mente humana um tipo especial de mente, por ser linguística. Nesse sentido, Dennett argumenta que apenas os humanos são convocados para dar razões para suas ações, o que cria a impressão da consciência, isto é, uma mente privada e restrita ao sujeito.

Passemos, agora, ao segundo grupo de perguntas. A primeira pergunta era essa: Na perspectiva interpretivista, é possível estabelecer uma relação entre os estados mentais e o mundo objetivo? Ora, aqui se pode perceber de imediato que esse é o fundamento do interpretivismo: os estados mentais existem porque foram construídos a partir de uma relação com o mundo objetivo. No caso de Davidson, é a relação de triangulação entre o sujeito, seus interlocutores e o mundo. Para Dennett, podemos regredir na escala evolutiva e procurar as

razões¹³⁶ porque temos estados mentais. Essas razões vinculam os estados cognitivos do organismo ao ambiente no qual evoluiu. Daí a tendência a que nossas crenças forneçam informações verdadeiras tanto acerca da realidade quanto sobre outras mentes.

Deve-se notar, porém, que estamos diante apenas de uma tendência à formação de crença verdadeiras. Um detalhe muito importante é que, conforme Dennett, nossos sistemas cognitivos são capazes de capturar aspectos da realidade. Isso ocorre porque fomos construídos por um processo imperfeito que relacionava nossas capacidades ao mundo conforme nossas necessidades reprodutivas, em situações contingentes e locais. Daí porque nem todos os aspectos da realidade objetiva estão disponíveis a nós e os aspectos que estão disponíveis são capturados por nós de modo imperfeito, mas isso não significa que estejamos impossibilitados de capturar a realidade objetiva. Ao contrário, nós a capturamos. Além disso, as atribuições de estados mentais que fazemos aos outros sujeitos são formas de capturar essa realidade¹³⁷.

Aqui, podemos fazer uma revisão das hipóteses que haviam sido colocadas quando a pesquisa se iniciava: A primeira hipótese era que ambos os autores respondessem afirmativamente à pergunta acerca da realidade dos estados mentais. Aparentemente, as pesquisas corroboram com essa hipótese. Inicialmente, havia uma dúvida em relação a se a posição de Dennett seria instrumentalista ou realista, mas a resposta é que, em Dennett, temos um realismo que parte do instrumentalismo. Em outras palavras, os estados mentais são reais porque são instrumentais (úteis e indispensáveis).

A segunda hipótese, contudo, não foi corroborada da mesma maneira. A hipótese era que só Davidson tivesse uma resposta afirmativa à questão acerca de se os estados mentais direcionavam o sujeito ao mundo objetivo e que Dennett não estaria preocupado com essa questão. Porém, ao longo das pesquisas foi possível perceber que a situação não era bem essa. Dennett está, sim, preocupado em mostrar que há um direcionamento dos nossos sistemas cognitivos ao mundo objetivo e isso se revela de modo mais contundente por meio do conceito de “affordances”, bastante explorado em seu último livro.

¹³⁶ Tratam-se, nesse caso, dos raciocínios de livre flutuação da natureza. Nós, animais linguísticos, podemos representá-los. Ver seção: 4.3.2.

¹³⁷ Davidson foi um autor que se projetava no tempo. Para ele, suas ideias seriam iguais às de Spinoza se tivesse vivido no passado (DAVIDSON, 1993/2005d). Provavelmente, não haveria contradição entre esse pensamento de Davidson e a ideia de Dennett de ancorar a realidade do mental no processo evolutivo se Davidson tivesse vivido no futuro, embora não possamos afirmar isso com certeza.

CONSIDERAÇÕES FINAIS: O PONTO DE VISTA DE TERCEIRA PESSOA

Ao longo da tese, apresentaram-se argumentos em defesa da possibilidade teórica de que o mental seja abordado por um ponto de vista de terceira pessoa. Tal abordagem não deve implicar em uma posição eliminativista acerca dos estados mentais, nem em uma posição epifenomenalista. Em outras palavras, os estados mentais estão aí e possuem relações de causa e efeito com o mundo externo, seja por terem sido provocados pelas situações ocorridas nele, seja por provocarem reações comportamentais nos sujeitos aos quais se atribuem esses estados mentais.

Para que a abordagem acima descrita seja possível, será necessário abrir mão de se considerar o mental a partir de um ponto de vista estritamente subjetivo, pois isso gera alguns problemas muito misteriosos para serem tratados. Em última instância, quando se toma o mental como algo restrito ao sujeito, nada se pode dizer sobre ele, apenas o sujeito poderá desfrutá-lo sozinho e ensimesmado. Afinal, as próprias palavras se lançam no mundo, levando o sujeito a compartilhar seus estados cognitivos com os outros, a partir de um idioma comum. Esse idioma foi aprendido a partir das relações que o sujeito manteve com o mundo físico e social, ao longo de sua história de vida. Nesse sentido, o mental se torna manifesto na linguagem.

Argumentou-se, a partir de uma ideia lançada por Sellars (1956/1997) que o mental pode ser compreendido como uma construção feita de fora para dentro. Sellars busca uma explicação para o significado que seja independente de um vocabulário mentalista. Para ele, a linguagem se desenvolve a partir do reforço a determinados comportamentos dos indivíduos dentro de uma comunidade linguística (ROSEMBERG, 2005). Esse mesmo tipo de explicação é ampliado para englobar os pensamentos. Os pensamentos, conforme o autor, são eventos intencionais internos, mas que possuem um papel explicativo frente aos comportamentos públicos. Eles se constituem a partir dos itens linguísticos, publicamente aprendidos.

Para esclarecer essa ideia, Sellars constrói o Mito de Jones, no qual pede para imaginarmos uma comunidade primitiva cujo vocabulário refere-se apenas a itens espaço-temporais de acesso público. Com o passar do tempo, a linguagem dessa comunidade é incrementada. Primeiro, ela desenvolve noções semânticas como significado, referência e

verdade, para que os membros da comunidade possam esclarecer seus discursos. Posteriormente, a linguagem desenvolve um vocabulário acerca de termos teóricos não observáveis, para formar um quadro geral do mundo observável. Note-se, aqui, que tudo partiu das relações entre os itens publicamente disponíveis para que depois os novos estágios da linguagem pudessem emergir.

Surge nessa comunidade um gênio chamado Jones que observa o comportamento inteligente de seus colegas, mesmo quando não estão usando a linguagem pública. Jones propõe uma hipótese segundo a qual “as enunciações públicas nada mais são do que a culminância de um processo que se inicia com certos episódios internos” (SELLARS, 1956/1997, p. 186, tradução nossa). Jones passa a oferecer explicações para os comportamentos dos sujeitos a partir de seus estados internos. A partir daí, o herói Jones ensina à comunidade como fazer esse tipo de explicação do comportamento, ampliando o vocabulário e os recursos explicativos daquela linguagem para incluir referências a estados mentais internos.

Assim, Sellars mostra como é possível transferir para o mundo interno de um sujeito um universo semântico que, inicialmente, poderia ser explicado a partir de padrões de comportamento e uma linguagem publicamente aprendida para se referir aos aspectos publicamente disponíveis do mundo. O discurso do sujeito sobre seus estados internos, mesmo sendo fruto de um acesso privilegiado, provém de uma construção que se inicia na realidade intersubjetiva e depois se transfere para o mundo íntimo. De acordo com Maroldi (2014), Sellars mostra como uma comunidade passa de uma linguagem fundamentada unicamente em situações publicamente compartilhadas (incluindo os comportamentos) para um vocabulário acerca dos estados mentais internos.

A posição interpretivista de Donald Davidson e de Daniel Dennett pode ser considerada uma derivação dessa ideia de Sellars que busca compreender os estados externos como algo que permite ao sujeito falar sobre si mesmo. Sellars mostra a possibilidade teórica de que essa explicação tenha se constituído em um estágio posterior do desenvolvimento da linguagem, a qual se inicia pela via dos comportamentos e objetos publicamente acessíveis. Com isso, abre-se uma via nova pela qual o mental pode ser abordado, um caminho a partir de fora.

Dennett oferece uma compreensão do mental como algo que surgiu gradualmente ao longo da história evolutiva. Em nossa espécie, o mental se tornou mais complexo, um mundo interior sobre o qual podemos falar, explicando as razões de nossos comportamentos. Porém, esse mundo subjetivo é, para Dennett, uma ilusão linguisticamente constituída. O autor

ênfatiza a importância disso para a constituição de um sujeito de responsabilidade moral. Nós somos capazes de explicar as razões pelas quais agimos, e por isso respondemos moralmente por nossas ações. Nesse sentido, o universo mental subjetivo só se torna possível com o advento da linguagem complexa e se torna indispensável para as organizações humanas em sociedades, pois constitui o fundamento de nossa responsabilidade moral.

Para Davidson, agir de acordo com razões é algo que nos caracteriza como seres humanos. As razões, para ele, são definidas a partir de atitudes proposicionais do agente que favorecem determinadas ações. Por isso, as razões porque agimos é o que nos torna animais racionais. “À luz de uma razão primária, uma ação é revelada como coerente com certos traços - de longo ou curto prazos - característicos ou não do agente, e o agente é mostrado em seu papel de Animal Racional” (DAVIDSON, 1963, p. 690, tradução nossa). Há vários desejos e crenças envolvidos na tomada de decisão comportamental, não é um conjunto preciso ou bem delimitado de atitudes proposicionais e poderia ser explicado de formas ligeiramente diferentes em diferentes momentos, mas é esse conjunto difuso que explica os comportamentos dos agentes.

Davidson foca nos estados mentais proposicionais e restringe a capacidade de pensar apenas ao humanos, o que gera uma lacuna entre suas duas formas de triangulação, a linguística e a pré-linguística. O trabalho de Tomasello com primatas, nesse sentido, pode ajudar a diminuir essa lacuna entre os dois tipos de triangulação.

A posição que busca estudar os estados mentais a partir do ponto de vista de terceira pessoa, de que se tratou ao longo da tese, se denomina “interpretivismo”. De acordo com essa posição, os estados mentais do sujeito dependem, de algum modo, da forma como seu comportamento pode ser interpretado. Algumas criaturas - sejam elas biológicas ou artificiais - se comportam de tal maneira que não é possível lhes dar sentido a não ser por meio da atribuição de estados mentais. Isso não é uma negação de que tais criaturas estejam sujeitas ao mesmo fluxo causal ao qual tudo se submete. Ao contrário, se fosse feito um mapeamento detalhado dos processos físicos subjacentes ao seu comportamento, seria possível explicá-lo, mas não seria possível compreender as razões para as ações. Essas criaturas, então, precisam ser compreendidas a partir da atribuição de estados mentais. São chamadas por Dennett de “sistemas intencionais”.

Interpretivismo não é eliminativismo

Alguns poderiam compreender o interpretivismo como uma posição puramente instrumentalista, isto é, fingimos que determinadas criaturas possuem mentes para podermos dar sentido ao seu comportamento, mas isso não significa que a criatura em questão possua uma mente de fato. A razão porque atribuiríamos mentes a certos sistemas, segundo essa leitura do interpretivismo, seria puramente prática, mas no fundo a posição consistiria em uma negação da existência de mentes dentro dos sujeitos. Apenas adotamos o método de tratá-los *como se* eles tivessem mentes, por razões meramente práticas.

Searle (2004), por exemplo, apresenta a posição de Davidson com um tipo de materialismo eliminativo. Materialismo eliminativo é uma posição segundo a qual os estados mentais não existem e por isso o vocabulário mentalista deve ser substituído por um vocabulário científico (da neurociência, por exemplo). Em parte, as raízes dessa posição estão nos mesmos autores que abrem o caminho para o interpretivismo, isto é, Sellars e Quine (RAMSEY, 2016). Ao apresentar a possibilidade de que o mental seja compreendido como termo teórico, Sellars abre a perspectiva do eliminativismo. Quine, por sua vez, chega a sugerir que o vocabulário mentalista seja substituído por um vocabulário mais baseado em aspectos observáveis da fisiologia. Na leitura de Searle, quando Davidson nega o fisicalismo de tipo, acaba por eliminar o mental.

Searle também compreende o pensamento de Dennett como eliminativista:

Uma variante da visão eliminativista é o que nós poderíamos chamar de “interpretativismo”¹³⁸. A ideia aqui é que as atribuições de intencionalidade são sempre formas de interpretação feitas por observadores externos. Uma versão extrema dessa visão é a concepção de Daniel Dennett de que nós algumas vezes adotamos a “postura intencional” e que nós não deveríamos pensar nas pessoas como tendo literalmente crenças e desejos, mas sim que essa é uma postura útil para se adotar acerca delas com o intuito de prever o seu comportamento (SEARLE, 2004, p. 115, aspas internas do original, tradução nossa).

Essa também é a leitura que Heil (1998) faz do interpretivismo. Ele compreende essa posição como um tipo de instrumentalismo que nega a realidade dos estados mentais. O interpretivismo, na leitura de Heil, compreende os estados mentais como uma espécie de convenção cultural feita para explicar e prever os comportamentos dos membros daquela comunidade. A analogia é com os sistemas de coordenadas, como os meridianos e paralelos imaginários que são sobrepostos ao mapa do mundo para que possamos nos localizar melhor

¹³⁸ O termo usado por Searle é “interpretativism”. Aqui, o termo “interpretivismo” está sendo usado, mas estamos ambos nos referindo à posição defendida por Dennett e Davidson.

frente a ele. Esses meridianos foram traçados por meio de simples convenções culturais. Embora tenham sido frutos de estudos geográficos sistemáticos e detalhados, não são algo que esteja no mundo. Assim seria, segundo Heil, a compreensão que os autores interpretivistas possuem dos estados mentais.

Foi dito que Davidson faz uma analogia das atribuições de estados mentais com sistemas de medição, mas o ponto da analogia, no caso deste autor, é outro. O que Davidson pretende mostrar é que há mais de uma maneira possível de atribuir estados mentais a um sujeito para explicar seu comportamento, assim como algo pode ser medido por sistemas diversos. Por exemplo, podemos efetuar a medição de um quadro em polegadas ou em centímetros e não há como decidir qual seria a maneira correta de efetuar a medição. Da mesma forma, segundo Davidson, é possível que dois sistemas de atribuição de estados mentais a um sujeito sejam igualmente explicativos em relação àquele sujeito, e nesse caso não seria possível definir qual o sistema correto. A analogia com os sistemas de medição acaba levando Davidson a um tipo de realismo, contrariamente à leitura de Heil. Isso porque, embora não seja possível estabelecer o sistema de medição apropriado, há algo fixo subjacente às diferentes medições, ou seja, aquilo que está sendo medido. Nessa mesma direção, Davidson pensa que há algo fixo sob as diferentes atribuições de estados mentais.

A analogia com os sistemas de medição, portanto, é usada por Davidson para buscar uma medida que ele considera apropriada para a indeterminação da interpretação, uma ideia herdada de Quine tanto por ele quanto por Dennett. Dennett pensa, porém, que, diante de uma situação de indeterminação, não haverá algo fixo subjacente. Por isso, Dennett não é tão entusiasta da analogia com os sistemas de medição. Mas isso também não leva Dennett ao eliminativismo. As atribuições de estados mentais, para ele, são tão indispensáveis em nossas relações quotidianas que não podemos abrir mão delas.

Essa leitura de Searle e de Heil acerca do interpretivismo, portanto, não é exata. O que acontece é que os autores como Dennett e Davidson não privilegiam o ponto de vista subjetivo em seus estudos acerca do mental. O ponto de partida desses autores é invertido: eles pensam que é de fora para dentro que se poderá formar uma compreensão da mente, e não de dentro para fora. Isso não é um tipo de eliminativismo nem é uma compreensão do mental como uma convenção cultural feita para se explicar e prever o comportamento. A questão, para esses autores, é que o mental existe para que o sujeito possa lidar com situações que o relacionam com o mundo físico e social. É a partir dessas relações que se torna possível uma compreensão do mental mais objetiva. Se há aspectos puramente subjetivos do mental, serão deixados para depois, mas a esperança é de que nada reste encarcerado no sujeito. Eles

partem dos aspectos relacionais dos estados mentais, e não do ponto de vista subjetivo. Ocorre que alguns autores pensam que isso não pode ser feito, que a única via para a explicação do mental deve ser de dentro para fora. Por isso, consideram o interpretivismo como um tipo de eliminativismo.

De fato, para muitos filósofos, o problema central não foi como proporcionar uma abordagem científica da experiência consciente, mas como penetrar o “véu de percepção” e ir “daqui” para o “mundo externo”, e as *Meditações* de Descartes foram a exploração inaugural desse modo de pensar.

O preço que você paga para seguir o conselho de Searle é que você toma todo o seu *phenomena*, os eventos e coisas que têm que ser explicados pela sua teoria, por um canal projetado não para investigação científica, mas para um uso prático, rápido e sujo na aspereza e queda da vida pressionada pelo tempo (DENNETT, 2017b, p. 365, tradução nossa).

De acordo com Dennett, nós temos a impressão de que alguns de nossos estados mentais são completamente subjetivos. Essa impressão nos é gerada para que possamos lidar com as situações que enfrentamos quotidianamente em tempo real, quais sejam, as situações de interação com o mundo, em especial, com as outras pessoas. Porém, tentar explicar o mental a partir dessa impressão subjetiva acaba por nos encarcerar em um caminho sem saída. Por isso, é melhor tentar explicar a impressão de subjetividade a partir dos aspectos compartilháveis do mental. Como nós temos a linguagem objetiva para expressar nossos estados subjetivos, então esse seria o melhor caminho e os estados mentais proposicionais seriam os mais profícuos.

Interpretivismo também não é epifenomenalismo

Há um nítido empenho tanto de Davidson quanto de Dennett para negar o epifenomenalismo. Eles pretendem mostrar que os estados mentais são provocados pelas situações do mundo exterior, assim como possuem um papel causal nas decisões comportamentais. Ao mesmo tempo, é importante para esses autores defender uma irreduzibilidade do vocabulário mental, garantindo que nós agimos conforme razões. Portanto, a causação deve ser mantida sem que se adote uma redução do vocabulário mental a um vocabulário da física. Provavelmente, esse é o cerne do interpretivismo: para um sistema ser considerado dotado de estados mentais, ele precisa ser interpretável como tal. Essa interpretação precisa ser feita a partir de um vocabulário mental. As propriedades semânticas podem ser eliminadas quando se pretende compreender o cérebro, mas não quando se tenta

compreender a mente. Em outras palavras, há processos causais que só podem ser compreendidos a partir de suas propriedades semânticas.

Essa ideia é apresentada por Davidson por intermédio do monismo anômalo. Segundo essa posição, o pensamento possui uma eficácia causal no mundo material e isso é compatível com as leis que regem a natureza. Na perspectiva de Davidson, um evento mental é um evento físico, daí o seu monismo. Porém, há mais de uma maneira de se descrever os eventos. Então, para um evento descrito como mental, há também uma descrição puramente física. Essa descrição, embora esteja de acordo com as leis estritas da física, não é capaz de capturar determinados aspectos relevantes para uma descrição apropriada daquele evento. Os agentes se comportam em conformidade com razões. Para que o comportamento do agente possa ter sentido, é preciso um apelo a essas razões, que remetem ao estados mentais daquele agente.

Na concepção de Davidson, não é possível a construção de leis que reduzam o mental ao físico. Por isso, embora seja um fisicalista, rejeita o fisicalismo de tipo¹³⁹. Cada evento mental específico é um evento físico, mas não há correlações legiformes entre tipos de eventos mentais e tipos de eventos físicos. Davidson considera que as razões são causas das ações, sendo característico de qualquer ação humana que ela possa ser explicada por meio de referências às intenções do agente (SHIMMENTI, 2012). Nesse sentido, não abre mão do papel causal dos estados mentais. Porém, as relações causais estão sempre sujeitas a uma lei geral, ao passo que as relações entre razões e causas não se submetem a leis gerais, são relações específicas e contextuais (NANNINI, 1999). Então, para que as razões se manifestem como causas, devidamente sujeitas a leis estritas, é necessário dar a elas uma descrição física. Isso é possível, para Davidson, porque os eventos mentais são eventos físicos.

A leitura que Kim (1989) faz do monismo anômalo é que essa posição de Davidson não é capaz de conferir poderes causais aos estados mentais. Para Kim, a eficácia causal dos estados mentais depende da existência de leis gerais relacionando tipos de estados mentais a tipos de estados físicos, caso contrário, estaremos diante de um epifenomenalismo. Se é preciso uma descrição física do fenômeno para que se possa dar conta de seus poderes causais, na leitura de Kim, o mental não possui qualquer eficácia causal. Toda a causalidade do mundo físico poderia ser descrita, no fim das contas, sem qualquer referência aos estados mentais.

Porém, essa crítica de Kim parece partir de um pressuposto dualista, que de algum modo compreende descrições de eventos mentais como algo muito diferente de descrições de

¹³⁹ O fisicalismo de exemplares (“token physicalism”), adotado por Davidson, preconiza que todas as coisas particulares do mundo são coisas físicas. Assim, cada estado mental particular é um estado físico específico. O fisicalismo de tipo correlaciona tipos gerais de estados mentais com tipos de estados físicos. Por exemplo: “dor é ativação da fibra C”, no sentido em que qualquer dor será uma ativação dessa fibra.

eventos físicos. Para Davidson, eventos são particulares e eventos mentais são eventos físicos particulares. Não são as descrições específicas dos eventos que possuem um papel causal, na leitura de Davidson, são os próprios eventos que estão relacionados causalmente, independente da maneira como são descritos.

Dennett (2013) também precisa se defender da acusação de epifenomenalismo, e seu caminho também é a adoção de um fisicalismo de exemplares particulares. O autor elabora um experimento de pensamento detalhado que serve ao mesmo tempo para esclarecer sua posição e para mostrar que não é epifenomenalista: o experimento das “duas caixas pretas”. Ao mesmo tempo em que mostra a possibilidade de as relações micro causais subjacentes a qualquer evento particular serem mapeadas por um cientista, mostra também que nem sempre esse mapeamento poderá levar a uma compreensão daquele evento. Isso porque, em muitos casos, apenas um apelo a noções semânticas, como significado, referência ou verdade, poderão dar conta de explicar o que está acontecendo. Esse experimento de pensamento resume as ideias de Dennett acerca das relações causais, mostrando diversos detalhes importantes.

O experimento de pensamento das duas caixas pretas foi proposto inicialmente em 1992, como uma resposta de Dennett (1992b) a Kim. O artigo “Real Patterns” de Dennett (1991a) foi considerado uma proposta epifenomenalista, algo que Dennett não estaria disposto a aceitar. No livro de 2013, o experimento de pensamento é apresentado de modo bastante detalhado.

Dennett pede ao leitor para supor que um grupo de cientistas tenha encontrado duas grandes caixas pretas conectadas por um fio: caixa A e caixa B. A caixa A possuía dois botões: α e β . A caixa B possuía três lâmpadas: uma vermelha, uma verde e uma amarela. Os cientistas observaram que a luz vermelha piscava uma vez sempre que eles pressionavam o botão α e a luz verde, quando eles pressionavam o botão β . A luz amarela quase nunca se acendia. Fizeram inúmeras observações em condições diversas, sempre com o mesmo resultado, até que estabeleceram a existência de uma regularidade:

- O botão α causa o piscar uma vez da luz vermelha e o botão β causa o piscar uma vez da luz verde.

A partir daí, eles passaram a acreditar que havia algo passando a causalidade pelo fio que conectava as duas caixas, mas o que seria? Eles testaram a intensidade do pulso que passava no fio, sua duração e velocidade, mas não conseguiram descobrir nada muito simples. Após inúmeras hipóteses serem falsificadas, perceberam que sempre que algum botão na caixa A era apertado, uma sequência complexa de muitos pulsos e pausas passava pelo fio.

Porém, o padrão da transmissão era diferente a cada vez, a mesma sequência de pulsos e pausas nunca se repetia, e por isso os cientistas continuaram intrigados, sem conseguir compreender qual era a relação causal entre os botões e as luzes.

Os cientistas decidiram, então, abrir a máquina B, onde encontraram um supercomputador digital. Perceberam que as sequências de pulsos e pausas eram códigos computacionais binários. Perceberam também que, em cada caso específico, era possível acompanhar os processos ocorridos no computador e que provocavam o acendimento da luz vermelha ou da luz verde, mas esses processos internos também eram específicos para cada caso. Assim, eles não foram capazes de construir nenhuma generalização, apenas de compreender as relações micro causais ocorridas nas situações particulares.

Eles resolvem lançar para a máquina B suas próprias sequências de pulsos e pausas, mas isso acaba por aumentar o mistério, pois a luz amarela piscava nesses casos. Eles retomam sequências que haviam sido lançadas pela caixa A anteriormente e obtêm o mesmo resultado anteriormente obtido. Ou seja, sempre que as mesmas sequências eram lançadas, os mesmos resultados eram obtidos, a luz vermelha se acendia quando apertavam o botão α e a luz verde, quando apertavam o botão β . Isso era verificado nos casos particulares, onde eles encontravam um perfeito determinismo, mas não conseguiam encontrar nenhuma forma de generalizar as relações entre as caixas. Abrem a caixa A, onde também encontram um supercomputador o qual nunca manda a mesma sequência para a caixa B, mas sempre mantém o padrão de fazer acender a luz vermelha quando seu botão α é apertado e a luz verde quando se aperta o botão β (a não ser em casos muito raros, como eles observam após uma grande quantidade de testes). Além do mais, se os cientistas reproduzem as sequências de pulsos e pausas da caixa A, obtêm os mesmos resultados que seriam obtidos se as sequências fossem enviadas pela própria caixa.

Eis que surgem os programadores das caixas para explicar o que está acontecendo. Cada um deles tinha desenvolvido independentemente o projeto de criar uma máquina de verdades. Eles equiparam suas máquinas com diversas verdades, além de regras de inferência para que suas máquinas pudessem construir novas verdades. Esses programadores eram de países diferentes, e por isso suas máquinas continham frases verdadeiras em línguas diferentes. Eles buscaram equipar suas máquinas com verdades que lhes pareciam de interesse geral. Por exemplo, listas de ganhadores do Prêmio Nobel, dados sobre personalidades famosas, dados de história geral dos países, temas de direito e relações internacionais, descobertas científicas diversas. Evitaram lançar temas de interesse restrito, como a cor da roupa que o programador usou quando saiu da maternidade onde nasceu seu filho, ou quantos passos em média ele

costuma dar entre sua cama e a pia da cozinha. Por isso, as máquinas acabaram sendo equipadas com um fundo amplo de verdades compartilhadas.

Uma vez que haviam sido equipadas com muitas verdades sobre diversos assuntos e com regras de inferência, essas máquinas eram muito boas em testar frases contra seu pano de fundo de verdades e dizer se eram verdadeiras ou falsas, em seus respectivos idiomas. Fora isso, os programadores passaram a não ver muita utilidade para suas máquinas e deixaram o projeto inacabado. Eles decidiram, então, unir as duas máquinas por um fio e programá-las para fazer traduções de suas frases para um idioma comum e codificado em linguagem de máquina, criando uma espécie de máquina semântica.

A nova máquina funciona assim: se o botão α é acionado, A escolhe aleatoriamente uma de suas “crenças”¹⁴⁰ e a traduz para o idioma comum e codificado, enviando o resultado a B. B traduz a frase recebida e a testa contra sua própria base de dados. Assim, se aquela frase recebida for uma “crença” de B, ele acende a luz vermelha. Caso contrário, acende a luz verde. Considerando que ambas as caixas foram equipadas com uma grande quantidade de verdades, há uma tendência a que elas compartilhem das mesmas crenças. Se A envia para B algo que acredita ser verdade, B sinaliza a sua concordância. Se envia algo que acredita ser falso, B também concorda, sinalizando com a luz verde. Quando os cientistas se intrometiam na transmissão, o resultado era uma frase mal formada, ao que a máquina B sinalizava acendendo a luz amarela.

Os programadores explicam, então, qual é a propriedade causal misteriosa que os cientistas procuram: é a propriedade de ser uma frase verdadeira no idioma comum entre A e B que faz a luz vermelha acender e a propriedade de ser uma frase falsa, faz acender a luz verde. Se houver dúvidas quanto ao fato de A e B realmente conhecerem a verdade ou a falsidade, pode-se fazer uma nova definição das propriedades causais em questão. A luz vermelha se acende quando há uma coincidência entre algo que A e B acreditam ser verdadeiro e a luz verde se acende quando coincide algo que elas acreditam ser falso. O ponto principal, aqui, é que a propriedade causal procurada pelos cientistas não poderá ser encontrada a menos que se recorram a termos semânticos. Por mais que se possam compreender os aspectos micro causais subjacentes nos casos específicos do funcionamento das duas caixas pretas, uma compreensão capaz de englobar uma generalização só pode ser feita em um nível de explicação mais elevado: atribuindo-se semanticidade ao sistema.

¹⁴⁰ Esse experimento de pensamento funcionará independentemente do leitor aceitar que a máquina tinha crenças ou não.

Alguns filósofos diriam que não há uma regularidade causal, dada a impossibilidade de se construir uma generalização nas sequências de pulsos e pausas transmitidas na máquina e que verdade e falsidade, por serem abstratas, não poderiam ter poderes causais. Dennett não pensa assim. Ele acredita que há uma regularidade causal, e os cientistas podem encontrá-la percorrendo o caminho micro causal exato que explica o resultado em cada caso particular. Aqui, Dennett está adotando um fisicalismo de casos particulares, assim como Davidson, e mostrando como o fisicalismo de tipo não convém quando estamos lidando com questões de nível alto, como o nível do discurso semântico.

Outro aspecto importante desse experimento de pensamento é mostrar a irreduzibilidade dos aspectos semânticos aos aspectos sintáticos (para quem alega que as máquinas são puramente sintáticas, Dennett propõe uma pequena alteração no experimento de pensamento: que as máquinas internas das caixas A e B sejam substituídas pelos seus respectivos programadores). O funcionamento do sistema composto pelas duas caixas pretas só pode ser compreendido pelo recurso às noções de verdade e falsidade, além das noções de significado e tradução. A única forma de explicar a regularidade entre A e B é ascender para o nível semântico, compreendendo as verdades como algo em que se crê e as frases emitidas como intencionais.

Muito esclarecedor também nesse experimento de pensamento é o fato de A e B compartilharem de quase todas as suas crenças. Isso é plausível, segundo Dennett, porque seus programadores, embora tenham vivido em países diferentes, habitam o mesmo mundo. Além disso, cada um deles procurou equipar o banco de dados de sua própria máquina com informações que pareciam de interesse geral, buscando criar uma espécie de enciclopédia. Nesse ponto do experimento de pensamento fica claro que, para Dennett, o fato de duas pessoas habitarem o mesmo mundo é suficiente para que tenham uma alta taxa de crenças comuns. Sem essa base de crenças compartilhadas, não seria possível compreender o funcionamento da máquina. Cada caixa contém a descrição de um mundo, e o mundo de ambas as caixas é o mesmo.

O ponto principal desse experimento de pensamento de Dennett é mostrar que não é possível abrir mão da postura intencional. Ou a postura intencional é adotada, ou a regularidade ali existente não poderá ser explicada. O discurso mental é indispensável e irreduzível.

Retomada de dois argumentos em favor do realismo moderado

Ao longo do presente trabalho, houve um esforço para argumentar que o interpretivismo adota um realismo brando em relação aos estados mentais do sujeito, em especial às atribuições de estados mentais. Em outras palavras, por um lado, as atribuições de estados mentais correspondem a algo, ainda que não seja algo dentro da cabeça do sujeito, mas algo sem o qual a comunicação não seria possível. Não é um realismo forte, pois os estados mentais não possuem existência concreta, mas também não é uma posição que descambe para o antirrealismo.

Há dois argumentos corroborando com essa tese, um de Davidson e o outro de Dennett, sendo que o argumento de Davidson é mais forte, e por isso também acaba carregando um pouco mais no seu realismo. Seu argumento é o da *triangulação*, que vincula mente, mundo e linguagem, de modo que eles não podem ser compreendidos isoladamente. O argumento de Dennett é de que, de modo imperfeito, local e contingente, os sistemas cognitivos dos animais foram constituídos para que eles pudessem lidar com as situações ambientais enfrentadas ao longo do processo evolutivo. Este argumento precisa ser encarado com muita parcimônia, dado o caráter local do processo evolutivo em que uma adaptação pode perder sua razão de ser frente às situações evolutivas encontradas pelo organismo e ainda assim permanecerem como característica daquela espécie.

Davidson apresenta dois tipos de triangulação. A primeira, chamada por ele de “primitiva”, não requer intencionalidade. “Envolve duas (ou, como sempre, mais que duas) criaturas reagindo à mesma cena, evento ou objeto, e correlacionando as reações das outras com os estímulos externos observados” (DAVIDSON, 1997/2001h, p. 292, tradução nossa). Isso é necessário, segundo o autor, para se estabelecerem condições necessárias ao pensamento e ao discurso, embora ainda não se estabeleçam as condições suficientes. A segunda forma de triangulação ocorre no contexto da aprendizagem da linguagem. O processo funciona pelo condicionamento do aprendiz a associar o comportamento do instrutor a aspectos do mundo que o aprendiz é capaz de discriminar. Assim se forma a camada básica da linguagem do aprendiz e todo o discurso que posteriormente ele aprende possui uma história que remete ao processo triangular de aprendizagem da linguagem¹⁴¹.

Triangulação é o que ocorre sempre que necessitamos exibir algo no mundo para determinar o significado do discurso de alguém (VERHEGGEN, 2013). Para que possa haver

¹⁴¹ A triangulação linguística também pode acontecer quando duas pessoas tentam se comunicar em idiomas diferentes.

linguagem, são requeridos os dois tipos de triangulação. Esse argumento fundamenta a concepção externista de Davidson segundo a qual os conteúdos de nossos discursos e pensamentos dependem em parte da nossa história de interações com o ambiente. Sem essas interações, o discurso de uma pessoa não teria sentido. Além disso o argumento da triangulação serve de suporte para a noção de objetividade. É necessário que as criaturas em interação entre si reconheçam a existência de um mundo independente das crenças de cada uma delas.

Uma vez que a triangulação é necessária tanto para a existência da linguagem quanto para os próprios pensamentos, então o ceticismo acerca do mundo exterior e o ceticismo acerca das outras mentes ficam sem espaço para serem cultivados. Em outras palavras, acatado o argumento da triangulação, esses tipos de ceticismo perdem sua razão de ser, pois eles partem de uma concepção de filosofia na qual se buscam definições isoladas para termos como “sujeito”, “mundo” e “linguagem”. Em Davidson, os termos não se definem isoladamente porque as coisas a serem definidas não existem isoladamente, isso não significa que elas não existam. Ao contrário, as existências de umas apoiam as das outras. Davidson rejeita qualquer tipo de empirismo que dependa da existência de intermediários epistêmicos entre o mundo e as crenças que temos acerca dele. Para o autor, muitas das nossas crenças estão diretamente relacionadas com o mundo objetivo.

Eu não duvido que a experiência seja essencial para o nosso conhecimento do mundo, ou que os sentidos servem de mediadores entre o mundo e nosso conhecimento sobre ele. Até onde me lembro, eu jamais questioneei a existência de sensações não conceituadas. O que eu critiquei foi a introdução de intermediários *epistêmicos* entre o mundo e nossas crenças acerca dele; por intermediários epistêmicos eu quero significar itens que são dados na experiência e que fornecem as razões últimas para nossa opinião sobre o ambiente (DAVIDSON, 2001k, p. 285, tradução nossa).

Embora as crenças estejam em contato direto com o mundo objetivo, isso não significa, obviamente, que o mundo seja dependente das crenças. As crenças são objetivas na medida em que sua verdade ou falsidade independem do sujeito. “Nossos outros pensamentos e atitudes, na medida em que têm conteúdo proposicional, também são objetivos” (DAVIDSON, 2001l, p. 1, , tradução nossa).

Davidson critica o empirismo que compreende como as bases de nosso conhecimento os dados dos sentidos, independente de se esses dados possuem ou não alguma conexão com o mundo objetivo. Esse tipo de empirismo leva ao subjetivismo (consequentemente ao ceticismo) que, na concepção do autor, só pode ser combatido por meio da adoção de uma perspectiva externista. O externismo compreende a ligação entre pensamento e mundo como

algo que reside na própria fundamentação do pensamento. No caso de Davidson (2001), seu externismo se fundamenta no processo de triangulação entre duas criaturas em comunicação entre si e com o mundo. Embora Davidson considere o empirismo de Quine subjetivista, por basear nossas crenças em algo privado e dado aos sentidos, percebe um aspecto externista também na sua tese da tradução radical. Isso porque a tradução radical mostra que o sentido dado por um falante a seu próprio discurso está intrinsecamente relacionado à interpretação competente daquele discurso (que depende da atribuição de significados e estados mentais). “Há, então, certa ambivalência na posição de Quine. Por um lado, a epistemologia é empirista, baseada na ‘evidência privada’, por outro lado, a teoria do significado, como implementada pelo processo de tradução radical, é basicamente social” (DAVIDSON, 2001, p. 11, tradução nossa).

Quanto a Dennett, é um autor ao qual também se atribuiu um realismo moderado. Dennett (1991a) apresenta-se como próximo a Davidson em relação ao status ontológico das atribuições de estados mentais, mas considera sua posição ontologicamente mais branda do que a de Davidson. Isso porque Dennett é mais radical em relação à indeterminação, ao passo que Davidson a compreende como a simples constatação de que há mais de uma maneira de se dizer a mesma coisa.

Dennett possui uma compreensão das atribuições de estados mentais como algo a tal ponto necessário em nossas vidas que não podemos viver sem elas. Por isso, elas são mais do que um termo teórico útil, elas derivam sua realidade da sua importância epistemológica. A comparação feita por Dennett é com os centros de gravidade. Eles são pontos reais na medida em que podem ser determinados com precisão. Porém, essa determinação pode ser inútil (como o centro de todas as meias desaparecidas de Dennett) ou extremamente útil, como o conjunto difuso de estados mentais que atribuímos às outras pessoas e a nós mesmos. O relevante, nesses casos, é a utilidade, mais do que a existência dos objetos abstratos.

O realismo de Dennett e de Davidson é moderado porque os estados mentais, na concepção deles, não são algo que se possa localizar dentro do sujeito. Ao contrário, eles são distinguíveis por meio do comportamento daquele sujeito, o que não significa que possam ser eliminados em um discurso científico. As atribuições de estados mentais são tão reais que sem elas não haveria linguagem e, como diz Davidson (2001), não haveria sequer nenhum tipo de ceticismo, nem mesmo o ceticismo acerca da realidade dos estados mentais.

Note-se, aqui, que esses autores não apenas consideram que as atribuições de estados mentais possuem uma referência no mundo intersubjetivo, como muitas de nossas outras crenças também possuem a característica de se referirem diretamente ao mundo objetivo.

Assim, nós acabamos por formar um corpo de crenças majoritariamente verdadeiras e majoritariamente coincidente com os corpos de crenças das outras pessoas que dividem conosco da mesma realidade. Ainda que eu discorde de um amigo que diz que a imortalidade da alma foi comprovada pelos experimentos desenvolvidos pela física quântica, o simples fato de que eu sou capaz de conversar com ele exige que tenhamos uma base massiva de crenças compartilhadas.

Ocorre que os estados mentais, para Dennett, são reais na medida em que são discerníveis por alguém (um intérprete, digamos). Da mesma forma, ele considera que os padrões na natureza são reais na medida em que são discerníveis. Porém, as diversas espécies de seres vivos possuem diferentes sistemas cognitivos, de modo que capturam padrões diferentes. Mas isso, para Dennett, não significa um relativismo. Os padrões são reais, eles estão na natureza e são diversos. Nós não podemos capturar todos, pois nossos sistemas cognitivos não permitem, mas podemos capturar alguns deles, os que são relevantes para as situações que enfrentamos em nossas vidas biológicas e culturais. Aqui se pode ver que, embora haja um realismo em Dennett, há também uma negação do essencialismo.

Desde que Sócrates foi pioneiro na busca de saber o que todos os *Fs* têm em comum, em virtude do qual eles são *Fs*, o ideal de limites claros, afiados, tem sido um dos princípios fundadores da filosofia. As formas de Platão geraram as essências de Aristóteles, que geraram uma série de maneiras de perguntar as condições necessárias e suficientes, que geraram tipos naturais, que geraram marcadores de diferença e outras formas de arrumar as bordas de todos os conjuntos de coisas no mundo.

Quando Darwin veio com a descoberta revolucionária de que os conjuntos dos seres vivos não eram eternos, com bordas duras, dentro ou fora das classes, mas sim populações históricas com limites difusos, ilhas historicamente conectadas a outras ilhas pelo desaparecimento dos canais, as principais reações dos filósofos eram ignorar esse fato difícil de negar ou tratá-lo como um desafio: Agora, como devemos estabelecer nossa teoria dos conjuntos das formas de biscoitos nessa porção vaga e errante da realidade? (DENNETT, 2017c, p. 9-10, tradução nossa)

Um ponto de fundamentação tanto para o anti-essencialismo quanto para o realismo brando de Dennett é a teoria da evolução darwinista. O autor possui uma leitura adaptacionista do processo evolutivo. Para ele, a compreensão desse processo passa pela compreensão das razões (sem racionalidade) da natureza, considerando, é claro, o caráter contingente e local dessas razões. Dennett argumenta, por exemplo, que o “verde” não pode ser considerado um tipo natural porque ele é um produto da seleção natural, no sentido em que a seleção nos equipou com o aparato necessário para que o verde apareça para nós. Assim, os limites do que se pode definir como verde são absolutamente difusos e imperfeitos. Além disso, não podemos estabelecer com precisão os limites de uma espécie viva, pois as espécies

mais recentes derivam das espécies mais ancestrais. Nós somos humanos, somos primatas, somos mamíferos e assim sucessivamente. Pertencemos ao mesmo tronco evolutivo de onde descendem as algas, os vírus e as baratas.

Porém, ele pensa que, se a vida de alguma criatura depende de ela perceber certos aspectos do ambiente em um determinado momento, a natureza irá equipar aquela criatura para que ela possa perceber esses aspectos do ambiente (caso contrário, a criatura morrerá sem deixar descendentes). Aqui se manifesta o adaptacionismo de Dennett. Assim, ainda que nossas percepções sejam imperfeitas, elas são percepções de aspectos reais de um ambiente que possui grandes porções independentes de nós. Daí a sua defesa de que nós e os demais seres vivos somos capazes de capturar aspectos da realidade a partir de nossos aparatos cognitivos.

Isso se amplia para as atribuições de estados mentais. Na concepção de Dennett, nossa psicologia de senso comum é inata e já estava presente no ancestral comum entre nós e os grandes símios. Dennett se apoia em pesquisas recentes nas ciências cognitivas para fundamentar essa tese. A natureza nos equipou com a capacidade de atribuímos intencionalidade a outras criaturas porque isso foi relevante para nós. Mas o fato de que nossas atribuições de atitudes proposicionais aos nossos pares costuma funcionar para a manutenção da interação indica sua força. Determinados sistemas não podem ser compreendidos sem que lhes atribuamos intencionalidade, eles são sistemas intencionais e se comportam conforme determinados padrões de racionalidade irreduzíveis a explicações de nível mais básico. Assim, na concepção de Dennett, os estados mentais dos outros fazem parte de nossa ontologia, isto é, dos aspectos do ambiente que nós somos capazes de capturar.

Os *qualia* foram eliminados?

Numa perspectiva interpretivista, os estados mentais são algo construído historicamente, a partir do contato dos organismos com o mundo. Duas grandes histórias podem ser contadas acerca dos estados mentais: a história evolutiva (filogenia) e a história de vida de cada ser dotado de intencionalidade (ontogenia). De acordo com a primeira, nossos ancestrais tiveram que lidar com uma série de situações ao longo de sua existência que os levaram a atribuir estados intencionais aos membros de sua espécie. Posteriormente, na espécie humana, surgiu a capacidade de comunicação linguística a qual nos levou a criar nossos próprios mundos subjetivos, a partir da necessidade de fornecer razões para nossas

ações. Note-se, porém, que nessa perspectiva a subjetividade não é restrita ao sujeito. Ela é um fruto da fricção do sujeito com o mundo.

Os *qualia*, considerados como estados mentais privados, subjetivos, intrínsecos e inefáveis são eliminados, sim, por Dennett. Seus argumentos mostram os paradoxos aos quais é conduzido quem adota esse tipo de perspectiva acerca do mental. Isso não é uma negação de que tenhamos impressões, sensações ou percepções do mundo, nem que nossas impressões sejam experiências do sujeito, nem que nossas experiências sensoriais possam nos enganar. O que se está negando é o subjetivismo, a ideia de que haja um mundo interior rico de informações e completamente inacessível a qualquer perspectiva de terceira pessoa.

Nós percebemos de forma altamente imperfeita aspectos do mundo objetivo (COHEN; DENNETT; KANWISHER, 2016, tradução nossa). Isso possui influência em nossas tomadas de decisão comportamental, mas também nos comportamentos automatizados. Uma parte dessas percepções é retida na memória e podemos ter acesso a esses conteúdos quando precisamos. Esse acesso pode ser linguístico ou não. Por mais que seja possível uma divergência entre nossas percepções e o mundo objetivo (por sermos míopes, por estarmos loucos, sob efeito de drogas ou doentes do cérebro ou mesmo pelas imperfeições próprias da evolução, ou pelo que quer que seja, nos enganamos com grande frequência), essas imperfeições não costumam ser massivas ou globais, sendo que a base de acertos supera os erros. Até mesmo essas divergências são produzidas a partir do nosso contato direto com o mundo. Os erros e enganos ainda são pequenos frente aos acertos. Provavelmente, o foco nas imperfeições tem gerado uma perspectiva subjetivista ao longo da história do pensamento filosófico, pois enfatiza as diferenças entre nossas percepções e o mundo objetivo. A perspectiva interpretivista reconhece essas diferenças, mas se estrutura sobre as relações diretas do sujeito com o mundo, quer dizer, pelas coincidências entre nossos sistemas cognitivos e o mundo.

Caridade e racionalidade

Um importante ponto em comum entre Davidson e Dennett em relação ao interpretivismo é que ambos estão fundados nos princípios de caridade e na pressuposição de racionalidade. Segundo Dennett (1987, p. 343), esses são dois princípios normativos, herdados do pensamento de Quine, que indicam quais atitudes proposicionais devem ser

atribuídas às criaturas no processo de interpretação. Como mencionado¹⁴², Miguens (2002, p. 24), apresenta a teoria dos sistemas intencionais de Dennett como uma “uma teoria normativa ou transcendental da mente. Entende-se por teoria normativa ou transcendental uma teoria segundo a qual os objetos são inteligíveis em virtude de um compromisso prévio quanto àquilo que eles podem ser.” A mente, para Dennett, só pode ser caracterizada a partir de um ponto de vista de quem a interpreta. Nesse sentido, também podemos atribuir a Davidson uma teoria normativa ou transcendental.

Em geral, as posições naturalistas costumam se opor às normativas em filosofia, no sentido em que as teorias naturalistas tentariam compreender como as coisas são por si mesmas, seriam abordagens descritivistas e as teorias normativas buscariam compreender as coisas a partir de como elas deveriam ser. O interpretivismo de Davidson e Dennett, ao contrário, são abordagens simultaneamente naturalistas e normativas porque estabelecem o ponto de vista a partir do qual a mente deve ser estudada, mas esse ponto de vista é descritivo, é o ponto de vista do intérprete.

Ao longo de sua trajetória filosófica, Dennett amplia sua concepção de consciência. Ela passa a ser o resultado de uma série de eventos ocorrendo simultaneamente e sendo editados no cérebro do sujeito. Esses eventos não se resumem a atitudes proposicionais, incluem outros tipos de estados mentais. O fato de não serem proposicionais não torna esses estados mentais inefáveis. Eles podem se manifestar comportamentalmente de outras formas, havendo inclusive a possibilidade de que sejam mais facilmente acessíveis a partir de um ponto de vista de terceira pessoa do que para o próprio sujeito. Com isso, a atribuição de estados mentais para a interpretação de outras entidades também se torna algo mais amplo. Embora nós tenhamos a capacidade de atribuir crenças e desejos a outras entidades, isso é algo que fazemos de modo natural, e não proposicional.

Essa ampliação tem uma vantagem: considerando que a atribuição de intencionalidade pode ter se desenvolvido por sua importância preditiva ao longo da cadeia evolutiva dos animais, então não faz sentido restringi-la apenas aos seres linguísticos. A linguagem precisou ter uma base para se desenvolver e, antes de haver linguagem, já existiam animais que se comunicavam. Por isso, é preciso incluir atribuição de intencionalidade não proposicional, e Dennett é capaz de fazer isso, ampliando a aplicação do interpretivismo não apenas para os seres linguísticos, mas para qualquer sistema cujo comportamento possa ser explicado a partir da postura intencional.

¹⁴² Ver: Seção 3.

Para concluir, vale enfatizar que a perspectiva interpretivista exige algumas inversões relativamente às perspectivas tradicionais do que seja o sujeito. A primeira inversão é buscar uma perspectiva de terceira pessoa para explicar o sujeito, não olhar o sujeito a partir de si mesmo. A segunda inversão é tentar compreender como a mente do sujeito pode ter sido construída a partir de sua necessidade de compreender as mentes dos outros. A terceira é pensar sobre como, a partir de processos simples, pode se ter construído uma linguagem complexa, uma cultura complexa, e só a partir daí, as mentes complexas teriam surgido, num processo sem projetista, de baixo para cima. Por último, ressalte-se que essa perspectiva parte da rejeição ao subjetivismo. Há argumentos que favorecem essa rejeição, mas isso exige um esforço para olhar a realidade a partir de um ângulo bastante diferente daquele com o qual já estamos habituados.

REFERÊNCIAS

- ABRANTES, P. A psicologia de senso comum em cenários para a evolução da mente humana. *Manuscrito*, Campinas, v. 29, n. 1, p. 185-257, jan./ jun. 2006. Disponível em: <http://www.unb.br/ih/fil/pcabrant/ artigos/Cenarios.pdf>. Acesso em: 10 jan. 2009.
- _____. Evolução humana: estudos filosóficos. *Aurora*, Curitiba, v. 25, n. 36, p. 75-105, jan./ jun. 2013.
- AMARAL, F. Causação mental: onde estivemos e onde estamos. *Psicologia: Teoria e Pesquisa*, Brasília, v. 17, n. 3, p. 224-235, set. 2001.
- BAGHRAMIAN, M; CARTER, J. Relativism. In: ZALTA, E. (Ed.). *The Stanford Encyclopedia of Philosophy*. Winter, 2016 Edition. Disponível em: <https://plato.stanford.edu/entries/relativism/>. Acesso em 08 abr. 2016.
- BEISECKER, D. Whither Normativity? *Studia Philosophica Estonica*, p. 1-9, out. 2015. Disponível em: <https://www.spe.ut.ee/ojs/index.php/spe/article/view/222/112>. Acesso em 24 fev. 2018.
- BERKOWITZ, S.; LANEY, C.; MORRIS, E.; GARRY, M.; LOFTUS, E. Pluto Behaving Badly: False Beliefs and Their Consequences. *The American Journal of Psychology*. v. 21, n. 4, p. 643-660, winter, 2008,
- BLACKMORE, S. *The meme machine*. Oxford: Oxford University Press, 1999.
- _____. There is no Stream of Consciousness. *Journal of Consciousness Studies*, v. 9, n. 5-6, p. 17-28, maio/ jun. 2002.
- BLOCK, N. Inverted Earth. *Philosophical Perspectives: Action Theory and Philosophy of Mind*, v.4 , p. 53-74, 1990. Disponível em: <http://www.jstor.org/stable/2214187>. Acesso em: 05 mar. 08.
- _____. On a confusion about a function of consciousness. *Behavioral and Brain Sciences*, v. 18, n. 2, p. 227-87, 1995.
- BRANDON, R. Study Guide. In: SELLARS, W. *Empiricism and the philosophy of Mind*. Cambridge; London: Harvard University Press, 1997.

BYRNE, A. Inverted qualia. In: *Stanford Encyclopedia of Philosophy*. In: ZALTA, E. (Ed.) Winter, 2006 Edition. Disponível em: <http://plato.stanford.edu/entries/qualia-inverted/>.

Acesso em: 05 fev. 2008.

CALL, J.; TOMASELLO, M. Use of social information in the problem solving of orangutangus (*Pongo pigmaeus*) and human children (*Homo sapiens*). *Journal of Comparative Psychology*, v. 21, n.3, p. 308-20, jul. 1995.

CAMPBELL, D. Blind variation and selective retention in creative thought as in other knowledge processes. *Psychological Review*, v. 67, n. 6, p. 380-400, 1960.

CARO, M. Davidson in focus. In: _____ (Ed.). *Interpretations and causes: New Perspectives on Donald Davidson's Philosophy*. [Berlin?]: Springer-Science, 1999. p. 1-29.

CAVELL, M. Triangulação: a natureza social do pensamento. In: *Becoming a subject. Reflections on Philosophy and Psychoanalysis*. Oxford: Clarendon Press, 2006. [Tradução de Ronai Pires Rocha, 2012]. Disponível em:

https://www.academia.edu/2020517/Marcia_Cavell_-_Triangula%C3%A7%C3%A3o_a_natureza_social_do_pensamento. Acesso em 3 jun. 2017.

CHALMERS, D. *The conscious Mind*. Oxford: Oxford University Press, 1996.

_____. *The Meta-Problem of Consciousness*. Disponível em:

<https://philpapers.org/archive/CHATMO-32.pdf>. Acesso em 28 fev. 2018. [Rascunho inicial]

CHISHOLM, R. *Perciving*. Ithaca: Cornell University Press, 1957.

CHURCHLAND, P. *Matéria e Consciência*. São Paulo: Editora UNESP, 2004. [Tradução de Maria Clara Cescato. Publicado Originalmente em 1984.]

COHEN, M.; DENNETT, D.; KANWISHER, N. What is the Bandwidth of Perceptual Experience? *Trends in Cognitive Science*, v. 20, n. 6, p. 324-335, maio 2016. Disponível em:

http://ase.tufts.edu/cogstud/dennett/papers/Cohen_Dennett_Kanwisher_2016_TICS.pdf.

Acesso em 28 jun. 2018.

CRANE, T. Intentionality. In: CRAIG, E. (Ed.) *Routledge Encyclopedia of Philosophy*. New York: Routledge, 1998. p. 816-821.

DAVIDSON, D. Actions, Reasons and Causes. *The Journal of Philosophy*, v. 16, n. 23, p. 685-700, nov. 1963.

_____. Theories of Meaning and Learnable Languages. In: _____. *Inquiries into truth & interpretation*. Oxford; New York: Oxford University Press, 1984a. p. 3-15 [Artigo publicado originalmente em 1965.]

_____. Mental Events. In: _____. *Filosofia de la psicologia*. Barcelona: Anthropos, 1994b. [Artigo publicado originalmente em 1970. Tradução de Miguel Candel.]

_____. Radical interpretation. In: _____. *Inquiries into truth & interpretation*. Oxford; New York: Oxford University Press, 1984b. p. 125-139. [Artigo publicado originalmente em 1973.]

_____. On the very idea of a conceptual scheme. In: _____. *Inquiries into truth & interpretation*. Oxford; New York: Oxford University Press, 1984c. p. 141-154. [Artigo publicado originalmente em 1974.]

_____. Belief and the Basis of Meaning. In: _____. *Inquiries into truth & interpretation*. Oxford; New York: Oxford University Press, 1984d. p. 141-154. [Artigo publicado originalmente em 1975.]

_____. Thought and Talk. In: _____. *Inquiries into truth & interpretation*. Oxford; New York: Oxford University Press, 1984e. p. 155-170. [Artigo publicado originalmente em 1975.]

_____. Reality Without Reference. In: _____. *Inquiries into truth & interpretation*. Oxford; New York: Oxford University Press, 1984f. p. 215-226. [Artigo publicado originalmente em 1977.]

_____. The Inscrutability of Reference. In: _____. *Inquiries into truth & interpretation*. Oxford; New York: Oxford University Press, 1984g. p. 227-241. [Artigo publicado originalmente em 1979.]

_____. The first person authority. In: _____. *Subjective, Intersubjective, Objective*. Oxford: Clarendon Press, 2001g. p. 3-14. [Artigo publicado originalmente em 1984.]

_____. A nice derragent of epitaphs. In: LEPORE, E. (Ed.). *Truth and interpretation: Perspectives on the philosophy of Donald Davidson*. Cambridge: Blackwell, 1986. p. 251-265.

_____. Knowing One's Own Mind. In: _____. *Subjective, Intersubjective, Objective*. Oxford: Clarendon Press, 2001a. p.15-38 [Artigo publicado originalmente em 1987.]

- _____. A Coherence Theory of Truth and Knowledge. *Subjective, Intersubjective, Objective*. Oxford: Clarendon Press, 2001b. p. 137-157. [Artigo publicado originalmente em 1983. Reflexões posteriores publicadas em 1987.]
- _____. The myth of the subjective. In: _____. *Subjective, Intersubjective, Objective*. Oxford: Clarendon Press, 2001j. p. 39-52. [Artigo publicado originalmente em 1988.]
- _____. What is present to the mind? In: _____. *Subjective, Intersubjective, Objective*. Oxford: Clarendon Press, 2001i. p. 53-68. [Artigo publicado originalmente em 1989.]
- _____. A estrutura e o conteúdo da verdade. In: _____. *Ensaaios sobre a verdade*. São Paulo: Editora Unimarco, 2002a. p. 47-107. [Organizado por Paulo Ghiraldelli Jr.; Pedro Bendassolli e Waldomiro José da Silva Filho. Tradução de Paulo Ghiraldello Jr. e Pedro Bendassolli. Artigo publicado originalmente em 1990.]
- _____. Epistemology Externalized. In: _____. *Subjective, Intersubjective, Objective*. Oxford: Clarendon Press, 2001e. p. 193-204. [Artigo publicado originalmente em 1990.]
- _____. Three varieties of knowledge. In: _____. *Subjective, Intersubjective, Objective*. Oxford: Clarendon Press, 2001f. p. 205-220. [Artigo publicado originalmente em 1991.]
- _____. The Second Person. *Subjective, Intersubjective, Objective*. Oxford: Clarendon Press, 2001c. p. 107-121. [Artigo publicado originalmente em 1992.]
- _____. Spinoza's Causal Theory of the Affects. In: _____. *Truth, Language and History*. Oxford: Clarendon Press; Oxford, 2005d. p.295-314. [Publicado originalmente em 1993]
- _____. Thinking causes. In: _____. *Truth, Language and History*. Oxford: Clarendon Press; Oxford, 2005b. p.185-200. [Publicado originalmente em 1993]
- _____. Radical Interpretation Interpreted. *Philosophical Perspectives*, [s.l.] v. 8, Logic and Language [s. n.], 1994a. p. 121-128. Disponível em: <http://www.jstor.org/stable/2214166>. Acesso em: 13 mar. 2017.
- _____. Davidson, Donald. In: GUTTENPLAN, Samuel. *A Companion to the Philosophy of Mind*. Oxford: Blackwell, 1995.
- _____. Pursuit of the concept of truth. _____. *Truth, Language and History*. Oxford: Clarendon Press; Oxford, 2005c. p.63-80. [Publicado originalmente em 1995]

_____. A tolice de tentar definir a verdade. In: _____. *Ensaio sobre a verdade*. São Paulo: Unimarco Editora, 2002b. p. 109-128 [Organizado por Paulo Ghiraldelli Jr., Pedro F. Bendassolli e Waldomiro José da Silva Filho. Tradução de Pedro Fernando Bendassolli. Artigo publicado originalmente em 1996.]

_____. Indeterminism and Antirealism. In: _____. *Subjective, Intersubjective, Objective*. Oxford: Clarendon Press, 2001h. p. 69-84. [Artigo publicado originalmente em 1997.]

_____. The Emergence of Thought. In: _____. *Subjective, Intersubjective, Objective*. Oxford: Clarendon Press, 2001d. p. 123-134. [Artigo publicado originalmente em 1997.]

_____. Appendix: Replies to Rorty, Stroud, McDowell, and Pereda. In: _____. *Truth, Language and History*. Oxford: Clarendon Press; Oxford, 2005a. p. 315-327. [Publicado originalmente em 1998]

_____. Interpretation: Hard in Theory, easy in practice. In: _____ (Ed.). *Interpretations and causes: New Perspectives on Donald Davidson's Philosophy*. Berlin (?): Springer-Science, 1999. p. 31-44.

_____. Verdade Reabilitada. In: _____. *Ensaio sobre a verdade*. São Paulo: Unimarco Editora, 2002c. p. 129-144. [Organizado por Paulo Ghiraldelli Jr., Pedro F. Bendassolli e Waldomiro José da Silva Filho. Tradução de Paulo Ghiraldelli. Artigo publicado originalmente em 2000.]

_____. Comments on Karlov Vary Papers. In: KOTATKO, P.; PAGIN, P.; SEGAL, G. *Interpreting Davidson*. California: CSLI Publications, 2001k. p. 285-307.

_____. Externalisms. In: KOTATKO, P.; PAGIN, P.; SEGAL, G. *Interpreting Davidson*. California: CSLI Publications, 2001l. p. 1-16..

_____. *Subjective, Intersubjective, Objective*. Oxford: Clarendon Press, 2001.

DAWKINS, R. *The Selfish Gene*. Oxford: Oxford University Press, 1976.

_____. *O relojoeiro cego: a teoria da evolução contra o desígnio divino*. São Paulo: Companhia das Letras, 2001. [Tradução de Laura Teixeira Motta. Publicado originalmente em 1986]

_____. *A escalada do monte improvável*. São Paulo: Companhia das letras, 1998. [Tradução de Suzana Sturlini Couto. Publicado Originalmente em 1996.]

_____. *The Ancestor's Tale: A Pilgrimage to the Dawn of Evolution*. Boston: Houghton Mifflin, 2004.

DENNETT, D. *Content and consciousness*. London and New York: Routledge, 1969.

_____. Intentional Systems. *The Journal of Philosophy*. v. 68, n. 4, p. 87-106, fev. 1971.

Disponível em: <http://www.jstor.org/stable/2025382>. Acesso em 10 jun. 2017.

_____. Conditions of Personhood. In: Rorty, A. O. (Ed.). *The Identities of Persons*.

California: University of California Press, 1976.

_____. Como mudar sua mente. In: _____. *Brainstorms: Ensaio Filosófico Sobre a Mente e a Psicologia*. São Paulo: UNESP, 2006a. p. 391-402. [Tradução de Luiz Henrique de Araújo Dutra. Publicado originalmente em 1978.]

_____. Three Kinds of Intentional Psychology . In: HEALEY, R. (Ed.) *Reduction, Time and Reality*. Cambridge: Cambridge University Press, 1981. p. 37-61.

_____. How to Study Human Consciousness Empirically or Nothing Comes to Mind.

Synthese. n. 52, v. 2, p. 159-180, 1982. Disponível em:

<https://dl.tufts.edu/bookreader/tufts:ddennett-1982.00008#page/1/mode/1up>. Acesso em 16 abr. 2017.

_____. *The Intentional Instance*. Cambridge; Massachusetts: The MIT Press, 1987.

_____. Quining qualia. In: MARCEL, A.; BISIACH, E. (Eds.).

Consciousness in Modern Science. Oxford: Oxford University Press, 1988. p. 42-77.

_____. The Interpretation of Texts, People and Other Artifacts. *Philosophy and*

Phenomenological Research. v. 50, p. 177-194, Supplement, autumn, 1990. Disponível em:

<http://www.jstor.org/stable/2108038>. Acesso em: 02 mar. 2018.

_____. Real Patterns. *The journal of Philosophy*. v. 88, n. 1, p. 27-51, jan. 1991a.

_____. *Conscious Explained*. London: Penguin Books, 1991b.

_____. The Self as a Center of Narrative Gravity. In: KESSELL, F.; COLE, P.; JOHNSON, D. (Eds.) *Self and Consciousness: Multiple Perspectives*. Hillsdale, NJ: Erlbaum, 1992.

_____. The Unimagined Preposterousness of Zombies. *Journal of Consciousness Studies*, v. 2, n. 4, p. 322-6, 1995.

_____. *Tipos de Mentes: rumo a uma compreensão da consciência*. Rio de Janeiro: Rocco, 1997. [Tradução de Alexandre Tort; revisão técnica de Marcus Pinto. Publicado originalmente em 1996.]

_____. *A perigosa ideia de Darwin: a evolução e os significados da vida*. Rio de Janeiro: Rocco, 1998a. [Publicado originalmente em 1995. Tradução de Talita M. Rodrigues.]

_____. *Brainchildren: essays on designing minds*. Cambridge; Massachusetts: The MIT Press, 1998b.

_____. *The Fantasy of First-Person Science*. Third Draft. Mar. 2001. Disponível em: <https://ase.tufts.edu/cogstud/dennett/papers/chalmersdeb3dft.htm>. Acesso em 18 abr. 2018.

_____. The diversity of Darwinian Explanations. In: _____. *Freedom Evolves*. London: Penguin Books, 2003a. p. 181-185.

_____. Explaining the magic of consciousness. *Journal of Cultural and Evolutionary Psychology*. v.1, n. 1, p. 7-19, 2003b. Disponível em: <https://ase.tufts.edu/cogstud/dennett/papers/explainingmagic.pdf>. Acesso em 17 mar. 2018.

_____. Who's on first? Heterophenomenology Explained. *Journal of Consciousness Studies*, v. 10, 2003c. Disponível em: <https://ase.tufts.edu/cogstud/dennett/papers/JCSarticle.pdf>. Acesso em 10 abr. 2018.

_____. How much is that in real money? In: *Sweet Dreams*. Massachusetts: The MIT Press, 2005a.

_____. *Sweet Dreams*. Massachusetts: The MIT Press, 2005b.

_____. Thank Godness! Edge: The Third Culture. 3 nov. 2006b. Disponível em: https://www.edge.org/conversation/daniel_c_dennett-thank-goodness. Acesso em 16 mar. 2018.

_____. Intentional Systems Theory. McLAUGHLIN, B.; BECKERMANN, A.; WALTER, S. *The Oxford Handbook of Philosophy of Mind*. Oxford: Oxford University Press, 2009. p. 339-349.

_____. *Intuition Pumps and Other Tools for Thinking*. New York; London: Norton & Company, 2013.

_____. Folk Psychology. *Intuition Pumps and Other Tools for Thinking*. New York; London: Norton & Company, 2013a. p. 73-76.

_____. An Older Brother Living in Cleveland. *Intuition Pumps and Other Tools for Thinking*. New York; London: Norton & Company, 2013b. p. 65-67.

_____. A Thing About Redheads. *Intuition Pumps and Other Tools for Thinking*. New York; London: Norton & Company, 2013c. p. 153-156.

_____. Illusionism as the Obvious Default Theory of Consciousness. *Journal of Consciousness Studies*. v.23, n. 11-12, p. 65-72, dez. 2016.

_____. I begrudge every hour I have do spend worrying about politics. *The Guardian*. 12 fev. 2017a. [Entrevistado por Carole Cadwalladr.] Disponível em: <https://www.theguardian.com/science/2017/feb/12/daniel-dennett-politics-bacteria-bach-back-dawkins-trump-interview>. Acesso em 17 mar. 2018.

_____. *From Bacteria to Bach and Back. The Evolution of Minds*. New York: W. W. Norton & Company, 2017b.

_____. Darwin and the Overdue Demise of Essentialism. In: SMITH, D. (Ed.). *How Did Biology Shapes Philosophy: New Foundations for Naturalism*. Cambridge: Cambridge University Press, 2017c.

DROZ, G. *Os mitos platônicos*. Brasília: Editora UnB, 1997. [Tradução de Maria Auxiliadora Ribeiro Kneipp. Original de 1992.]

EL-HANI, C. N.; PIHLSTRÖM. Realismo, Pragmatismo e Emergência. In: SMITH, P.; SILVA FILHO, W. (Orgs.) *Significado, verdade e interpretação: Davidson e a Filosofia*. São Paulo: Loyola, 2005. p. 199-238.

FAGUNDES, J. O. A. *A consciência vista de fora: a perspectiva de Dennett*. 2009. 134f. Dissertação (Mestrado em Filosofia) – Universidade de Brasília, Brasília, 2009. (Orientador: Paulo Cesar Coelho Abrantes).

_____. Davidson e o Obscurecimento do ceticismo por meio da linguagem. SÃO PAULO, I.; VINICIUS, M. (Eds.) *Revista Sísifo*, Feira de Santana, v.1, n.1, p. 18-26, maio 2015. Disponível em: http://www.revistasisifo.com/p/edicoes-anteriores_15.html. Acesso em 30 mar. 2017.

FODOR, J.; LEPORE, E. Is radical interpretation possible? *Philosophical Perspectives*. v. 8. Logic and Language, p. 101-119, 1994. Disponível em: <http://www.jstor.org/stable/2214165>. Acesso em 12 mar. 2017.

- FRANKISH, K. Editorial Introduction. _____. (Ed.) *Illusionism. A special double issue of Journal of consciousness studies. Journal of Consciousness Studies*. v. 23, n. 11-12, p.1-2, dez. 2016.
- FRECHÉTTE, G. A Tese de Brentano (Revisitada). *Revista Guaicará de Filosofia*. Guarapuava, v. 32, n. 2, p. 106-132, 2016.
- FREGE, G. *Investigações Lógicas*. Porto Alegre: EDIPUCRS, 2002. [Organização, tradução e notas de Paulo Alcoforado.]
- GIBSON, J. J. The Theory of Affordances. In: *The Ecological Approach to Visual Perception*. Boston: Houghton Mifflin, 1979, p. 127-137.
- GLOCK, H. *Dicionário Wittgenstein*. Rio de Janeiro: Jorge Zahar Editor, 1996. [Tradução de Helena Martins. Publicado originalmente em 1996.]
- GODFREY-SMITH, P. Conditions for Evolution by Natural Selection. *Journal of Philosophy*, v. 104, n. 10, p. 489-516, out. 2007.
- GOMES, H. *Facebook desligou robô que abandonou inglês e criou linguagem própria? Não é verdade!* Disponível em:
<https://g1.globo.com/e-ou-nao-e/noticia/facebook-desligou-robo-que-abandonou-ingles-e-criou-linguagem-propria-nao-e-verdade.ghtml>. Acesso em 25 ago 2017.
- GÓMEZ-TORRENTE, M. Alfred Tarski. ZALTA, E. (Ed.). *The Stanford Encyclopedia of Philosophy*. Spring 2015 Edition. Disponível em:
<https://plato.stanford.edu/archives/spr2015/entries/tarski/>. Acesso em 10 mar. 2017.
- GOULD, S. J.; LEWONTIN, R. C., The Spandrels of San Marco and the Panglossian Paradigm: A Critique of The Adaptationist Programme. *Proceedings of The Royal Society of London*, v. 205, n. 1161, p. 581-598, 1979.
- GRAHAM, G. Behaviorism. ZALTA, E. (Ed.) *The Stanford Encyclopedia of Philosophy*. Spring 2015 Edition. Disponível em:
<http://plato.stanford.edu/archives/spr2015/entries/behaviorism/>. Acesso em: 08 jul. 2015.
- HAACK, S. Is it true what they say about Tarski? *Philosophy*. Cambridge, v. 51, p. 323-336, abr. 1976.
- _____. Reflexions on Relativism: From Momentous Tautology to Seductive Contradiction. *Philosophical Perspectives*. v. 10, Metaphysics, p. 297-315, 1996.

- HEIL, J. *Filosofia da Mente: uma introdução contemporânea*. Lisboa: Instituto Piaget, 1998. [Tradução de Rui Pacheco.]
- HILTON, P. Willard van Orman Quine. ZALTA, E. (Ed.) *The Stanford Encyclopedia of Philosophy*. Spring 2018 Edition. Disponível em: <https://plato.stanford.edu/entries/quine/#IndTra>. Acesso em 10 jan. 2018.
- HIMMA, K. Design Arguments for the existence of God. *Internet Encyclopedia of Philosophy*. Disponível em: <http://www.iep.utm.edu/design/#SH1c>. Acesso em 26 set. 2017.
- JACOB, P. Intentionality. ZALTA, E. (Ed.) *The Stanford Encyclopedia of Philosophy* Winter 2014 Edition. Disponível em: <https://plato.stanford.edu/cgi-bin/encyclopedia/archinfo.cgi?entry=intentionality>. Acesso em 09 out. 2017.
- KIM, J. Can Supervenience and ‘Non-Strict Laws’ Save Anomalous Monism? In: Heil, J.; MELE, A. (Eds.) *Mental Causation*. Oxford: Clarendon Press, 1993. p. 19-23.
- KIRK, Robert. Sentience and Behavior. *Mind*, Oxford, v. 83, n. 239, p. 43-60, jan.1974.
- KRIEGEL, U. *The sources of intentionality*. Oxford: Oxford University Press, 2011.
- KRUPENYE, C.; KANO, F.; HIRATA, S.; CALL, J.; TOMASELLO, M. Great apes anticipate that other individuals will act according to false beliefs. *Science*, v. 354, n. 6308, p. 110-114, abr. 2016.
- KUHN, T. *A estrutura das revoluções científicas*. São Paulo: Editora Perspectiva, 1998. [Tradução de Beatriz Vianna Boeira e Nelson Boeira. Publicado originalmente em 1962.]
- LAU, J; DEUTSCH, M. Externalism about mental states. ZALTA, E. (Ed.) *The Stanford Encyclopedia of Philosophy*. Winter 2016 Edition. Disponível em: <https://plato.stanford.edu/cgi-bin/encyclopedia/archinfo.cgi?entry=content-externalism>. Acesso em 22 jun. 2018.
- LECLERC, A. Davidson’s externalism and swanpman’s troublesome biography. *Principia*. Florianópolis, v. 1-2, n. 9, p. 159-175, jun./dez. 2005.
- _____. Um tratamento contextualista do predicado de existência. *Filosofia Unisinos*. São Leopoldo, v. 7, n. 2, p. 113-125, maio/ago. 2006.
- _____. Intentionality and Continuity of experience. *Principia: an international journal of epistemology*. Florianópolis, v. 21, n. 2, p. 235-249, ago. 2017.

- LOCKE, J. *Ensaio sobre o entendimento humano*. São Paulo: Martins Fontes, 2012.
[Tradução, apresentação e notas de Pedro Paulo Garrido. Publicado originalmente em 1690.]
- LOFTUS, E. Creating false memories. *Scientific American*. v. 277, n. 3, set. 1997, p. 70-75.
Disponível em: <https://faculty.washington.edu/eloftus/Articles/sciam.htm>. Acesso em 22 mar. 2018.
- LOUX, M; SOLOMON, W. A inescrutabilidade e relatividade da referência segundo Quine. *Crítica na rede*. Disponível em: http://criticanarede.com/lin_quine.html. Acesso em 12 abr. 2017.
- MALPAS, J. Não renunciar ao mundo: Davidson e os fundamentos da crença. In: SMITH, P.; SILVA FILHO, W. (Orgs.) *Significado, verdade e interpretação: Davidson e a Filosofia*. São Paulo: Loyola, 2005. p. 51-65.
- _____. Donald Davidson. ZALTA, E. N. (Ed.). *The Stanford Encyclopedia of Philosophy*. Fall, 2015 Edition. Disponível em:
<https://plato.stanford.edu/archives/fall2015/entries/davidson/> Acesso em 12 mar. 2017.
- MAROLDI, M. Linguagem, Comportamento e Mente no Mito de Jones de Jones de Wilfrid Sellars. *Kínesis*, v. 6, n. 12, p. 89-105, dez. 2014. Disponível em:
https://www.marilia.unesp.br/Home/RevistasEletronicas/Kinesis/6_marcelomaroldi1.pdf.
Acesso em 20 jun. 2018.
- MCDOWELL. O dualismo esquema-conteúdo e empirismo. In: SMITH, P.; SILVA FILHO, W. (Orgs.) *Davidson e a Filosofia*. São Paulo: Loyola, 2005. p. 33-50.
- McGINN, C. Can We Solve The Mind-Body Problem? *Mind*, v. 93, n. 391, p. 349-366, jul. 1989.
- McKAY, R.; DENNETT, D. The Evolution of Misbelief. *Behavioral and Brain Sciences*. Cambridge, v. 32, p. 493-561, 2009.
- McLAUGHLIN, B. P. Epiphenomenalism. In: GUTTENPLAN, S. *A companion to the philosophy of mind*. Oxford: Blackwell, 1995.
- MIGUENS, S. *Uma teoria fisicalista do conteúdo e da consciência: D. Dennett e os debates da filosofia da mente*. Porto: Campo das Letras Editores, 2002.

_____. Conceito de crença, triangulações e atenção conjunta. In: MIGUENS, S.; MAURO, C. (Eds.) *Perspectives on rationality*. Porto: Universidade do Porto – Faculdade de Letras, 2006. p. 99-117. Disponível em: <http://ler.letras.up.pt/uploads/ficheiros/9971.pdf>. Acesso em: 20 jun. 2017.

MILLER, A. Ceticismo acerca do sentido (I) - Quine sobre a analiticidade e a tradução. In: _____. *Filosofia da Linguagem*. São Paulo: Paulus, 2010. [Tradução de Evandro Luis Gomes; Christian Marcel de Amorim; Perret Gentil Dit Maillard.]

NAGEL, T. Como é ser um morcego? *Cadernos de História e Filosofia da Ciência*, Campinas, v. 15, n. 1, p. 245-262, jan./jun. 2005. (Publicado originalmente em 1974. Tradução de Paulo Abrantes e Juliana de Orione).

_____. Is Consciousness an Illusion? *The New York Review of Books*. 9 mar. 2017.

Disponível em:

<http://www.nybooks.com/articles/2017/03/09/is-consciousness-an-illusion-dennett-evolution/?printpage=true>. Acesso em 4 jun. 2018.

OKASHA, S. Realismo e anti realismo. *Crítica na rede*. Disponível em:

<http://criticanarede.com/realismo.html>. Acesso em 1 maio 2017.

PARENT, T. Externalism and Self-knowledge. ZALTA, E. (Ed.) *The Stanford Encyclopedia of Philosophy*. Summer 2013 Edition. Disponível em:

<<http://plato.stanford.edu/archives/sum2013/entries/self-knowledge-externalism/>>. Acesso em 07 out. 2016.

PEREIRA, D. P. C. *Donald Davidson: subjetivo-objetivo*. O retorno ao cogito, 2016, 201 f. Dissertação (Mestrado em Filosofia) - Faculdade de Letras da Universidade do Porto, Porto, 2009 (Orientadora: Sofia Miguens).

POPPER, K. A Evolução e a árvore do conhecimento. *Conhecimento Objetivo*. Belo Horizonte: Itatiaia, 1992, p. 234-56. [Tradução de Milton Machado. Publicado originalmente em 1972.]

PREMACK, D.; WOODRUFF, G. Does the Chimpanzee have a theory of mind? *The Behavioral and Brain Sciences*. v.1, n. 4. 1978. p. 515-526.

PUTNAM, H. *Mathematics, Matter and Method: Philosophical Papers*, v. 1. Londres: Cambridge: 1975.

_____. *Reason, Truth and History*. Cambridge: Cambridge University Press, 1981.

QUINE, W. *Palavra e Objeto*. Petrópolis: Editora Vozes, 2010. [Tradução de Desidério Murcho. Publicado originalmente em 1960.]

_____. On the Reasons for the Indeterminacy of Translation. *Journal of Philosophy*, v.67, n. 06, p. 178-183, mar. 1970.

_____. *Pursuit of truth*. Cambridge; London: Harvard University Press, 1992.

RAMSEY, W. Eliminative Materialism. ZALTA, E. (Ed.). *The Stanford Encyclopedia of Philosophy*. Winter, 2016 Edition. Disponível em:

<https://plato.stanford.edu/cgi-bin/encyclopedia/archinfo.cgi?entry=materialism-eliminative>. Acesso em 02 mar. 2018.

RAVENSCROFT, I. Folk Psychology as a Theory. ZALTA, E. (Ed.). *The Stanford Encyclopedia of Philosophy*. Fall, 2016 Edition. Disponível em:

<https://plato.stanford.edu/cgi-bin/encyclopedia/archinfo.cgi?entry=folkpsych-theory>. Acesso em 10 out. 2016.

RICHERSON, P. J.; BOYD, R. *Not by Genes Alone: How Culture Transformed Human Evolution*. Chicago: The University of Chicago Press, 2005.

ROBB, D.; HEIL, J. Mental Causation. ZALTA, E. (Ed.) *The Stanford Encyclopedia of Philosophy*. Spring 2013 Edition. Disponível em:

<http://plato.stanford.edu/cgi-bin/encyclopedia/archinfo.cgi?entry=mental-causation>. Acesso em 27 abr. 2013.

RORTY, R. Introduction. In: SELLARS, W. *Empiricism and the philosophy of Mind*. Cambridge; London: Harvard University Press, 1997.

ROSEMBERG, J. Wilfrid Sellars. *Crítica na Rede*. 2005. Disponível em:

https://criticanarede.com/his_sellers.html#topo. Acesso em 20 jun. 2018. [Originalmente publicado em ZALTA, E. (Ed.) *Stanford Encyclopedia of Philosophy*. Summer 2002 Edition. Tradução de Eduardo Coutinho Lourenço de Lima.]

RUFFINO, M. Indexicais. BRANQUINHO, J.; SANTOS, R. (Eds.). *Compêndio em Linha de Problemas de Filosofia Analítica*. Lisboa: Centro de Filosofia da Universidade de Lisboa, 2014. 20 fls. Disponível em:

http://repositorio.ul.pt/bitstream/10451/16087/1/ruffino_indexicais_artigo.pdf. Acesso em 29 mar. 2017.

RYLE, G. *The Concept of Mind*. New York: Routledge, 2009. [Publicado originalmente em 1949]

SANCHES, M. *Donald Davidson: acerca da correspondência, coerência e cepticismo*. Coimbra: Angelus Novus, 2002.

SANTOS, R. Paradoxos Semânticos. BRANQUINHO, J.; SANTOS, R. (Eds.). *Compêndio em Linha de Problemas de Filosofia Analítica*. Lisboa: Centro de Filosofia da Universidade de Lisboa, 2014. 43 fls. Disponível em:

http://compendioemlinha.letras.ulisboa.pt/wp-content/uploads/2015/07/santos_r_2014_paradoxos-sem%C3%A2nticos_artigo.pdf. Acesso em 29 mar. 2017.

SEARLE, J. Intentionality and its place in nature. _____. *Consciousness and Language*. Cambridge: Cambridge University Press, 2002. p. 77-89.

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

SELLARS, W. *Empiricism and the philosophy of Mind*. Cambridge; London: Harvard University Press, 1997. [Publicado originalmente em 1956.]

SILVA FILHO, W. J. Davidson, a Metáfora e os Domínios do Literal. *Utopía y Praxis Latinoamericana*, v. 6, n. 15, p. 30-43, dez. 2001.

_____. Davidson, Sócrates e os instrumentos da Filosofia. In: GHIRALDELLI JR., P.; BENDASSOLLI, P. F.; SILVA FILHO, W. J. (Orgs.) *Donald Davidson: ensaios sobre a verdade*. São Paulo: Unimarco Editora, 2002. p. 149-163.

_____. Externalismo deflacionado. In: SMITH, P.; SILVA FILHO, W. (Orgs.) *Significado, verdade, interpretação: Davidson e a Filosofia*. São Paulo: Loyola, 2005. p. 155-167.

_____. Externalismo sem dogmas. *O que nos faz pensar*. Rio de Janeiro, v. 16, n. 21, p. 113-131, jun. 2007.

SMART, J. Quine's Philosophy of Science. In: DAVIDSON, D.; HINTIKKA, J. (Eds.). *Words and Objections*. Boston: D. Reil Publishing Company, 1969. p. 3-13.

SMITH, B. Internalism and Externalism in the Philosophy of Mind and Language. *Internet Encyclopedia of Philosophy*. Disponível em: <https://www.iep.utm.edu/int-ex-ml/>. Acesso em 29 jun. 2018.

- SMITH, P. Davidson para além do ceticismo. In: SMITH, P.; SILVA FILHO, W. (Orgs.) *Significado, verdade, interpretação: Davidson e a Filosofia*. São Paulo: Loyola, 2005. p. 127-154.
- SIMONS, D. J.; LEVIN, Daniel T. Change blindness. *Trends in cognitive Sciences*, v. 9, n. 7, p. 261-7, out. 1997
- STROUD, B. Interpretação radical e ceticismo filosófico. In: SMITH, P.; SILVA FILHO, W. (Orgs.) *Significado, verdade, interpretação: Davidson e a Filosofia*. São Paulo: Loyola, 2005. p. 103-126.
- TARSKI, A. The concept of truth in formalized languages. In: _____. (Ed.). *Logic, Semantics, Metamathematics*. Oxford: Oxford University Press, 1936. p. 152-278.
- TENNANT, N. Radical Interpretation, Logic and Conceptual Schemes. In: _____ (Ed.). *Interpretations and causes: New Perspectives on Donald Davidson's Philosophy*. [Berlin?]: Springer-Science, 1999. p. 71-93.
- TENNIE, C.; CALL, J.; TOMASELLO, M. Ratcheting up the ratchet: on the evolution of cumulative culture. *Philosophical Transactions of The Royal Society B: Biological Sciences*. n. 374. 2009. p. 2405-2415.
- THORNTON, T. Recent developments for naturalizing the mind. *Current Opinion in Psychiatry*, Walters Kluer, v. 24, n. 6, p. 502-506, ago. 2011.
- TOMASELLO, M. *A Natural History of Human Thinking*. Cambridge; Massachusetts: Harvard University Press, 2014.
- TYE, M. Qualia. ZALTA, E. (Ed.) *The Stanford Encyclopedia of Philosophy*. Winter 2017 edition. Disponível em <<https://plato.stanford.edu/cgi-bin/encyclopedia/archinfo.cgi?entry=qualia>> Acesso em 22 fev. 2018.
- ULRICH, F. A verdade sobre o lastro do bitcoin. *Infomoney*. 2014 Disponível em: <http://www.infomoney.com.br/blogs/cambio/moeda-na-era-digital/post/3206256/verdade-sobre-lastro-bitcoin>. Acesso em 02 abr. 2018.
- UMPIERES, R. Febre do Bitcoin vai se esgotar em um “crash espetacular”, diz ex-diretor de Orçamento da Casa Branca. *Infomoney*, São Paulo, 02 jan. 2018. Mercados/Bitcoin. Disponível em:

<http://www.infomoney.com.br/mercados/bitcoin/noticia/7183015/febre-bitcoin-vai-esgotar-crash-espetacular-diz-diretor-orcamento-casa>. Acesso em 17 jan. 2018.

VERHEGGEN, C. Triangulation. In: LEPORE, E.; LUDWIG, K. *A Companion to Donald Davidson*. Oxford: Wiley Blackwell, 2013. p. 456-471.

WIMMER, H.; PERNER, J. Beliefs about beliefs: representation and constraining function of wrong beliefs em young childrens's understanding of deception. *Cognition*. v. 13, n. 01, p. 103-128, 1983.

WITTGENSTEIN, L. *Investigações Filosóficas*. São Paulo: Abril Cultural, 1979. (Os Pensadores) [Tradução de José Carlos Bruni. Publicado originalmente em 1953.]

YUDKOWSKY, E. Zombies! Zombies? In: *Rationality: From AI to Zombies*. Berkeley: MIRI, 2015. Disponível em: <https://www.lesswrong.com/posts/fdEWWr8St59bXLbQr/zombies-zombies>. Acesso em 28 fev. 2018.