

O que é o externismo?*

Autora: Katalin Farkas

Tradução: Luiz Helvécio Marques Segundo

Revisão: André Abath

Resumo: O conteúdo da tese externista sobre a mente depende crucialmente do modo como definimos a distinção entre o “interno” e o “externo”. De acordo com a compreensão comum, o limite entre o interno e o externo é o crânio ou a pele do sujeito. Neste artigo argumento que a compreensão comum é inadequada, e que apenas a nova compreensão da distinção externo/interno sugerida por mim nos ajuda a entender o problema da compatibilidade do externismo e do acesso privilegiado.

Introdução

O externismo sobre o conteúdo cognitivo tem sido discutido há quarenta anos, e tornou-se quase que ortodoxia na filosofia da mente. Essa ortodoxia entende a tese externista bastante a grosso modo, sem que haja qualquer acordo unânime sobre a sua natureza precisa. Os detalhes de uma definição precisa não importam para certos propósitos, mas são importantes se queremos extrair conseqüências da doutrina; por exemplo, ao considerar a compatibilidade do externismo e do autoconhecimento. Esse debate chegou a um estado de confusão quase que desesperador, e a razão disso, penso, está em certa confusão sobre o que o externismo é. No que se segue, tentarei clarificar essa questão.

* Tradução de “What is Externalism?”, de Katalin Farkas, originalmente publicado em *Philosophical Studies*, 112: 187-208, 2003.

1- A linha divisória entre o externo e o interno

Diversas teses têm sido chamadas de “externistas”, mesmo na filosofia da mente. Não posso esperar discutir todas elas aqui, de modo que darei atenção àquilo que se pode chamar “externismo da Terra Gêmea”, a versão do externismo claramente motivada pelo tipo de argumentos da Terra Gêmea. Esse tipo de argumentos visa estabelecer – pela análise de exemplos concretos – que o seguinte é possível: que dois sujeitos teriam estados internos qualitativamente idênticos e ainda assim o conteúdo de (alguns) seus estados mentais seria diferente por causa de alguma diferença em seu ambiente externo.

¹ O protótipo desses argumentos é, com certeza, o argumento de Putnam em “The Meaning of ‘Meaning’”. Esses argumentos são usados para apoiar a tese do externismo, que pode ser formulada, por exemplo, como se segue:

- O conteúdo dos pensamentos de um sujeito depende, ou é individuado, por fatos *externos* ao sujeito; ou que
- O conteúdo dos pensamentos de um sujeito não sobrevém de seus estados *internos*; ou que
- Ter certos pensamentos pressupõe a existência ou a natureza particular das coisas que são *externas* ao sujeito.

Pode haver outras versões, mas todas elas concordam num ponto, a saber, traçam uma linha divisória entre o externo e interno ou entre algumas noções correlatas. Essas formulações certamente capturam pelo menos parte do que o externismo é, mas serão incompletas se não responderem a uma pergunta crucial: o que significam as expressões “externo” e “interno”? Como deveríamos traçar a linha entre o interno e o externo?

1 Alguns filósofos usam a estratégia da Terra Gêmea para definir o externismo. Veja por exemplo McLaughlin e Tye (1998), p. 285, e Davies (1998), p. 327. Jackson e Pettit traçam uma distinção entre conteúdo “estrito” e “amplo” em termos da noção de um *Doppelgänger*; veja Jackson e Pettit (1996), p. 220.

Há uma interpretação que parece ser aceitável em muitas discussões: a de que “externo” significa “*externo ao corpo ou à pele (ou ao cérebro) de um sujeito*”.² A tese externista afirma assim que o conteúdo dos pensamentos ou das frases de um sujeito depende dos fatos externos à sua pele. Essa concepção certamente tem apoio na formulação original de Putnam, a de que “os significados não estão na cabeça”. Tentarei, no entanto, mostrar que na verdade o ponto principal do externismo não é sobre os fatos individuais estarem dentro ou fora da pele.

Como eu disse, o meu interesse aqui é na versão do externismo claramente motivado pelos argumentos do tipo dos da Terra Gêmea, que apresentam dois sujeitos cujos estados internos são estipulados como sendo os mesmos. Isso sugere uma maneira de descobrir o que “externo” e “interno” significam. Temos de nos focar na *relação entre os Gêmeos*; que coisa é essa que eles compartilham e que, de acordo com o externista, não é suficiente para individuar o conteúdo de seus pensamentos? Se pudermos dizer o que é, então teremos uma apreensão do que o “interno” é; e tudo o que os Gêmeos não puderem compartilhar será “externo”.

Dada a suposição de que “interno” significa “dentro da pele” e “externo” “fora da pele”, a estratégia comum apresenta dois sujeitos cujos estados dentro da pele são (qualitativamente) fisicamente idênticos. Assim, a relação entre os Gêmeos é de *identidade na constituição física qualitativa*. No que se segue, argumento que a estipulação sobre a identidade na constituição física não é nem necessária e nem suficiente para o argumento do externista proceder. Isso significa que a interpretação de “interno” como “dentro da pele” é inadequada; a linha divisória entre o interno e o externo não deveria ser traçada na pele.

2 Os exemplos são muitos; veja por exemplo Davies (1998), p. 322; McLaughlin e Tye (1998), p. 285; McDonald (1998), p. 124, Boghossian (1997), p. 163. A maioria dos autores referidos neste artigo fazem suposições similares, veja McCulloch (1995), p. 189, Jackson e Pettit (1996), p. 220; Burge (1988), p. 650.

2- A estipulação da identidade na constituição física não é suficiente

Se um argumento é bem sucedido em mostrar que dois sujeitos cuja constituição física é idêntica mas que, não obstante, têm conteúdos mentais diferentes, essa conclusão excluirá diversas teorias da mente. Uma vez que estados cerebrais, estados funcionais e disposições comportamentais (numa certa interpretação) sobrevivem a estados corporais, o externismo fora da pele pode ser usado para refutar, por exemplo, a teoria da identidade, o funcionalismo, e (certas versões de) o comportamentalismo.³ Tal argumento, no entanto, não consegue lidar com versões dualistas do internismo. Para se opor ao dualismo num espírito externista, deveríamos apresentar dois sujeitos cujos estados mentais são idênticos *de acordo com o critério dualista*, e então argumentar que seus pensamentos são diferentes devido a alguma diferença em seu ambiente. Mas estipular a identidade qualitativa na constituição física dos sujeitos não será suficiente para assegurar a identidade dos estados mentais de acordo com a concepção dualista; estados de almas imateriais ou propriedades não físicas de estados mentais não precisam sobreviver a estados corporais.

Dada a ampla aceitação do materialismo, lidar com o dualismo talvez não seja considerado uma questão importante. O ponto, porém, não é tão polêmico, mas explicativo. A teoria cartesiana da mente é geralmente considerada como o arquétipo do internismo; mas a razão para se considerar Descartes um internista não pode ser a de que de acordo com sua teoria os estados mentais são individuados por estados corporais. Assim, a compreensão comum deixa inexplicado em que sentido o cartesianismo é uma teoria internista.⁴

De fato, o próprio Putnam pensava que seu externismo tinha de excluir o dualismo. Ao descrever o cenário da Terra Gêmea, ele diz “suponha que tenho um

3 Veja McCulloch (1995), p. 168.

4 Há filósofos que parecem reconhecer esse ponto e tentam definir o externismo de modo que seja aplicável a teorias imaterialistas. Veja por exemplo Burge (1986), pp. 118-119; Pettit (1986), pp. 17-18. Gabriel Segal distingue entre propriedades “intrínsecas” e “localmente sobrevenientes”, e define o internismo primariamente em termos do primeiro (Segal, 2000, p. 9 ff.). Penso, no entanto, que, por razões a serem esmiuçadas na próxima seção, a noção de “intrínseco” não é adequada para definir o internismo.

Doppelgänger na Terra Gêmea que é molécula a molécula ‘idêntico’ a mim [...] Se você for um dualista, então suponha que meu *Doppelgänger* tenha os mesmos pensamentos verbalizados que eu, que tem os mesmos dados dos sentidos, as mesmas disposições, etc.” (Putnam, 1975, p. 227). Em outro ponto ele diz que os Gêmeos têm as mesmas crenças, pensamentos, sentimentos, e assim por diante (*ibid.* p. 224). O problema com isso é familiar: o argumento original de Putnam visava apoiar o externismo sobre *significados*, e depois é que se estendeu ao externismo para o *conteúdo mental*. Se o objetivo do argumento da Terra Gêmea é mostrar que o conteúdo das crenças do Gêmeo é diferente, então você não pode apresentar o argumento dizendo que eles têm as mesmas crenças ou pensamentos. Assim, embora Putnam pudesse talvez se beneficiar da relação de “ter os mesmos pensamentos” para argumentar a favor do externismo semântico, a mesma formulação não pode ser usada num argumento a favor do externismo sobre o conteúdo mental.⁵

Não podemos caracterizar a relação entre os Gêmeos como “ter os mesmos pensamentos”. Mas se o argumento exigisse apenas que os Gêmeos fossem idênticos molécula a molécula, então o argumento não lidaria com o dualismo. Assim, a estipulação de identidade na constituição física não é suficiente para a formulação de um argumento externista geral.

3- A estipulação da identidade na constituição física não é necessária

Defensavelmente, pelo menos algumas doenças são tipos naturais: elas têm algumas propriedades superficiais (os sintomas) com base nas quais normalmente as identificamos, e alguma estrutura subjacente que é responsável pelas propriedades superficiais – por exemplo, certa inflamação causada por alguma bactéria. Podemos construir o seguinte caso da Terra Gêmea: suponha que a doença conhecida como “meningite” na Terra Gêmea tenha exatamente os mesmo sintomas e todos os efeitos que

5 Veja Burge (1982).

a meningite na Terra, e que na verdade não é causada pela bactéria *Meningococcus* (como na Terra), mas por uma bactéria diferente a qual chamaremos “XYZ”. Considere Oscar₁ na Terra, sofrendo de meningite antes de 1750, quando a bactéria causadora da meningite era desconhecida, e Oscar₂ na Terra Gêmea, que é tão similar a Oscar₁ quanto possível, exceto por ele estar sofrendo ao mesmo tempo da doença causada por XYZ. Parece, então, que o argumento poderia prosseguir do mesmo modo que no caso da Terra Gêmea: o pensamento de Oscar₁ “a meningite é uma doença perigosa” tem um conteúdo diferente do pensamento paralelo de Oscar₂, pois Oscar₁ está pensando sobre a meningite, e Oscar₂ sobre a doença causada por XYZ. Chegamos a uma conclusão muito similar em espírito à tese externista, mas os fatos individuadores nesse caso parecem estar dentro do corpo.⁶

Pode-se notar que o exemplo original da Terra Gêmea envolvendo “água” é de fato similar ao caso da meningite: como se tem chamado a atenção repetidamente, os Gêmeos do exemplo original não podem ser fisicamente idênticos, uma vez que o corpo humano contém uma quantidade significativa de água. Essa objeção geralmente não é considerada tão grave; a impressão geral foi a de que poderíamos encontrar um exemplo melhor sobre uma substância que não fosse encontrada no corpo humano, de modo que todos continuaram a usar o exemplo da água. Sugiro que a nossa disposição em negligenciar esse problema no argumento original é melhor explicada pelo fato de que o ponto principal do externismo não é o de os fatos individuadores estarem dentro ou fora do corpo. Na verdade, isso se torna ainda mais claro numa versão posterior do argumento que Putnam oferece: nessa versão, a água da Terra Gêmea contém 20 por cento de álcool, e a química do corpo das pessoas da Terra Gêmea é alterada de modo que elas possam reagir a essa mistura do mesmo modo que reagimos à água (Putnam, 1981, p.23). Esse argumento parece funcionar como um argumento a favor do externismo da Terra Gêmea, e, contudo, a condição da igualdade física interna é obviamente violada. O exemplo da meningite ajuda a tornar o ponto ainda mais claro,

6 Presumo aqui que o argumento causal é *independente* dos tipos burgianos de argumentos a favor do externismo social e dos argumentos baseados na divisão do trabalho lingüístico.

uma vez que o nosso estereótipo sobre a meningite é formado com base na sua ocorrência no corpo humano, ao passo que o mesmo não se dá com a água.

O que motiva a análise internista do argumento da água? Alguns filósofos recorrem a intuições simples, outros apóiam suas intuições com certa teoria dos termos para tipos naturais. Parece que qualquer que seja a motivação do exemplo original, ela também está no caso da meningite. Assim, se algum argumento a favor do externismo baseado nos termos para tipos naturais é digno de algo, o caso da meningite um exemplo tão bom quanto qualquer outra espécie de tipo natural. Ou, pondo de outro modo: se alguém dissesse que Oscar₁ e Oscar₂ se referiam a algo diferente por “meningite” por causa da diferença microscópica desconhecida de seus corpos, então teríamos um argumento perfeitamente bom a favor do externismo, em que a estipulação de que os Gêmeos seriam idênticos na constituição física não é necessária.

Soa-me óbvio que o ponto do argumento da meningite é exatamente o mesmo do argumento da água ou do alumínio, mas alguns talvez discordem. A objeção poderia correr assim: “ninguém nega que alguns fatos dentro da pele fazem diferença no conteúdo de nossos pensamentos e, portanto, não seria surpresa que sujeitos com constituições físicas diferentes tivessem pensamentos diferentes. O ponto dos argumentos da Terra Gêmea é que *ainda que* você estipulasse a identidade molécula a molécula, os pensamentos ainda assim poderiam diferir. Dado que temos essa tese *mais forte*, porque deveríamos nos preocupar com a tese mais fraca? De qualquer modo, o externismo é a tese de que os conteúdos mentais não sobrevivem aos estados corporais, de modo que a questão entre o internista e o externista é saber se *Gêmeos idênticos molécula a molécula* podem ter pensamentos diferentes ou não. Os protagonistas do exemplo da meningite *não são Gêmeos nesse sentido*, de modo que o que quer que digamos sobre seus pensamentos será irrelevante para a questão do externismo”.

Deixe-me oferecer a seguinte analogia para ilustrar o que me parece errado com essa objeção. Suponha que estipulamos que os nossos Gêmeos, Oscar₁ e Oscar₂, deveriam

ser idênticos molécula a molécula, e, ademais, que deveriam usar gravatas “idênticas”. Aplicamos então o argumento da Terra Gêmea comum e terminamos com a seguinte tese interessante: os conteúdos mentais não sobrevivem aos estados corporais mais os estados gravatais [*necktie states*]. De fato, interessante, dirão, mas você não poderia usar o argumento sem as gravatas? E a nossa resposta vem prontamente: se temos a tese mais forte sobre estados corporais *mais* os estados gravatais, por que deveríamos nos preocupar com a tese mais fraca? De qualquer modo, o externismo de gravata é sobre os Gêmeos que satisfazem das estipulações dadas; assim, o que quer que você diga sobre os Gêmeos que são *meramente* idênticos molécula a molécula, será irrelevante para os nossos propósitos.

Presumo que meus oponentes reconhecerão que defender o externismo de gravata é sem propósito, embora continuem ainda hesitantes. Eles agora admitirão que os Gêmeos não têm de ser idênticos molécula a molécula: por exemplo, Oscar₁ poderia ser uma polegada mais alto que Oscar₂, e ainda assim o argumento funcionaria. Mas embora pudéssemos todos concordar sobre casos particulares, eles insistirão que não há modo de *especificar o que conta em geral* como uma diferença relevante nos estados corporais – relevante no sentido em que concede ao externista o seu ponto – e o que não conta. Gravatas são arrancadas facilmente, mas partes do corpo não. Portanto, continua a objeção, o único modo lógico ou natural de traçar a linha divisória entre o externo e o interno é no corpo. Mas essa objeção funciona apenas se não houver de fato outro modo de traçar a linha entre o interno e o externo, e a tarefa deste artigo é mostrar o contrário.

4- Indistinguibilidade Subjetiva

Deixe-me sumarizar as duas seções precedentes. Primeiro, aceitei que o debate externismo/internismo é ortogonal ao debate materialismo/dualismo. Portanto, parece que uma tese externista geral seria efetiva também contra versões dualistas do internismo. Segundo, argumentei que o externismo – ou algo muito próximo disso –

pode surgir no que diz respeito aos fatos dentro do corpo. Se isso estiver correto, então podemos perguntar se é possível definir o externismo de modo que acomode esses dois pontos. Claramente que a nova definição tem de partir da ideia de que a pele é a linha divisória entre o interno e o externo. Admito que isso cria certa dificuldade: se a compreensão comum se baseia na concepção dentro da pele/fora da pele, então parece que simplesmente mudei o sujeito se propus uma modificação. Afinal, se vários filósofos explicitamente disserem – e dizem – o que querem dizer por externismo, então deveríamos tomar aquilo que dizem. Não quero disputar isso. Contudo, penso que é legítimo perguntar que outros motivos podem estar por trás da tese externista, e que podem ser trazidos à superfície considerando-se os pontos sobre o dualismo e o caso da meningite. Tentarei explicar como a minha proposta se sobrepõe à compreensão comum, e deixarei que o leitor decida entre elas. Argumentarei também que essa escolha terá conseqüências significativas no debate sobre a compatibilidade do externismo e do autoconhecimento.

Suponha então que pelo menos faz sentido perguntar se alguma compreensão do externismo pode acomodar os pontos levantados nas seções 2 e 3. O que precisamos então é de uma caracterização da relação entre os Gêmeos que (i) estabeleça uma identidade ou equivalência entre os Gêmeos em algum aspecto, (ii) torne os Gêmeos no caso da meningite equivalentes nesse aspecto; (iii) implique a igualdade “interna” de modo que seja aplicável a teorias dualistas. Como eu disse anteriormente, a caracterização dessa relação fornecerá a noção da linha divisória relevante entre o interno e o externo; o que quer que seja comum aos Gêmeos nesse aspecto é interno a um sujeito, e a conclusão externista é que isso é aquilo que é insuficiente para individuar o conteúdo mental.

Uma boa maneira para começar nossa busca pela caracterização propriamente dita é considerar modos possíveis de *danificar* os argumentos da Terra Gêmea a favor do externismo. É crucial para o experimento mental de Putnam que a água (H₂O) e a água gêmea (XYZ) sejam *indistinguíveis* para os agentes em circunstâncias perceptuais

normais (veja Putnam, 1975, p. 223). Se a água gêmea fosse azul e amarga em circunstâncias comuns, um internista poderia facilmente concordar que “água” significa algo diferente na Terra e na Terra Gêmea. Imagine que agora você vê um copo d’água, prova-o, e sacia sua sede. Sabemos que a situação da Terra e da Terra Gêmea é exatamente simétrica (eles chamam *o nosso* planeta de Terra Gêmea); ora, se *esta* fosse a Terra Gêmea com XYZ nela, o líquido chamado “água” pareceria, teria o mesmo cheiro e o mesmo gosto, e também saciaria a sua sede. (Nas versões em que a química do corpo varia da Terra para a Terra Gêmea, você deveria fazer uma troca contrafactual das químicas dos corpos).

O argumento da Terra Gêmea não funcionaria se a água e a água gêmea parecessem diferentes; e também não funcionaria se os estereótipos de água e água gêmea fossem diferentes. Suponha que os Gêmeos tivessem algum conhecimento de química; se fosse parte da concepção de água de Oscar₁ que é composta de H₂O, e parte da concepção de Oscar₂ que fosse composta de XYZ (no sentido de que se lhes perguntassem o que a água era, suas respostas incluíssem isso), então novamente um internista poderia facilmente concordar que “água” significa algo diferente na Terra e na Terra Gêmea. Assim, a fim de que o argumento a favor do externismo funcione, temos de excluir tais divergências; e é isso, acredito, o que Putnam tentou capturar ao dizer que Oscar₁ e Oscar₂ têm as mesmas crenças, pensamentos, etc. (Putnam, 1975, p. 224). Mas como eu disse anteriormente, não podemos usar a estipulação de que os Gêmeos têm os mesmos pensamentos ou crenças.

Como podemos capturar esses dois pontos? Recorramos a alguma ajuda. Burge é um dos poucos filósofos que discutem o problema de se formular uma posição individualista que seja aceitável às teorias não-materialistas. A sua sugestão inicial é que o individualismo é “a tese de que os fenômenos mentais qualitativos, fenomenológicos, fixam todos os estados mentais da pessoa, incluindo aqueles (como pensamentos, desejos, intenções) com características intencionais e representacionais” (Burge, 1986, p. 117). Essa caracterização parece se adequar ao aspecto do cenário da Terra Gêmea

que acabei de realçar: que a água e a água gêmea deveriam produzir as mesmas impressões e deveriam ser associados às mesmas características para os Gêmeos. Para usar a formulação de Burge, os Gêmeos deveriam ter os mesmos “fenômenos mentais qualitativos, fenomenológicos” ao experienciarem a água ou ao pensarem sobre ela.⁷

No entanto, Burge percebe um problema: essa caracterização pressupõe uma distinção bem compreendida entre aspectos *fenomenológicos* e *intencionais*, e Burge pensa que é bastante dúbio que Descartes e a tradição não-materialista (que supostamente sustentam essa versão de individualismo) estivessem cientes disso, ou tivessem aceito tal distinção. Felizmente, podemos capturar a mesma ideia sem precisar da maquinaria conceitual pesada que essa distinção envolve. A chave é: se você fosse (efetiva ou contrafactualmente) trocado com sua contraparte da Terra Gêmea, *as coisas pareceriam as mesmas*. Ser transportado para a Terra Gêmea (inconsciente, da noite para o dia) é diferente de ser transportado, digamos, para Marte. Nesse último caso, as coisas na manhã seguinte certamente pareceriam diferentes. Essa ideia parece-me constitutiva de um cenário da Terra Gêmea propriamente dito: que a sua situação é subjetivamente indistinguível da situação de seu *Doppelgänger*. Mas eu não gostaria de seguir Burge em pensar que a ideia pressupõe uma distinção entre o qualitativo e o intencional. Suponha que alguém não concorda com a compartimentalização dos fenômenos mentais em qualitativos e intencionais, e sustenta que todos os fenômenos mentais são intencionais. Seria bastante estranho se esse filósofo não pudesse conceitualizar o cenário da Terra Gêmea da maneira que fiz; se ele não pudesse entender a diferença entre ser transportado para a Terra Gêmea e ser transportado para Marte. Contudo, ele não explicaria isso dizendo que os fenômenos mentais qualitativos (como opostos aos intencionais) são os mesmos: ele não poderia, uma vez que não nega que tais coisas existam.

A noção de indistinguibilidade subjetiva é fundamental na compreensão da natureza da experiência humana, e é anterior à distinção qualitativo/intencional, ou ao

⁷ Presumo aqui, como diversos filósofos o fazem, que todo pensamento consciente tem caráter fenomenológico.

resultado do debate externismo/internismo. Para ilustrar esse último ponto, considere, por exemplo, a teoria disjuntiva da percepção. Os disjuntivistas sustentam que não há um único tipo de estado mental em que um sujeito esteja quando está tendo uma percepção verídica ou uma alucinação perfeita correspondente.⁸ Os disjuntivistas são externistas, pois o que faz a diferença entre os dois estados mentais em tipo é algo externo ao sujeito. Mas a fim de formular a teoria deveríamos ter uma compreensão do que é uma alucinação perfeita, e isso é dado precisamente em termos de indistinguibilidade subjetiva. Um cenário Gêmeo sugerido pelo disjuntivista poderia ser assim: você agora está vendo um copo d'água; mas você poderia estar alucinando, isto é, estar numa situação que é subjetivamente indistinguível desta, e contudo o copo não estar ali. Assim, os disjuntivistas aceitariam a caracterização da relação como indistinguibilidade subjetiva; no entanto, eles negariam que isso implique a identidade dos estados mentais. Penso que é nisso que os internistas e externistas se separam: um internista encontraria dificuldades em aceitar que algo que não faz diferença em como uma situação parece ao sujeito (a presença do copo no caso da percepção verídica, e sua falta no caso da alucinação perfeita correspondente) poderia fazer diferença aos seus estados mentais.⁹

A minha sugestão é esta então: a relação entre os Gêmeos é de indistinguibilidade subjetiva em suas situações. Repetindo, isso significa que se eles fossem trocados (efetiva ou contrafactualmente), as coisas pareceriam as mesmas. Essa é também a mesma relação existente entre uma percepção verídica e a alucinação correspondente. E a lição do argumento da Terra Gêmea a favor do externismo é que dois sujeitos que estão

8 Para uma concepção disjuntiva veja, por exemplo, McDowell (1982) ou McDowell (1986).

9 Novamente, uma vez que há muitas variedades de internismo, nem todos os internistas concordariam em apresentar a doutrina desse modo. Para um exemplo de um debate conduzido em termos similares, veja os comentários de John McDowell sobre Simon Blackburn no contexto de uma disputa sobre pensamentos dependentes de objetos. Blackburn descreve uma série de cenários do estilo da Terra Gêmea onde “tudo é o mesmo do ponto de vista do sujeito” e alega que “há uma categoria legítima de coisas que são as mesmas nesses casos” (Blackburn, 1984, p. 324). McDowell concorda quanto “a categoria incontroversamente legítima de coisas que são as mesmas através dos casos em que as coisas parecem ao sujeito”, mas nega que haja “algo” – um estado mental, por exemplo – que seja o mesmo através dessas situações (McDowell, 1986, p. 157).

em situações subjetivamente indistinguíveis poderiam estar em estados mentais diferentes. A rejeição do externismo da Terra Gêmea, por outro lado, é a negação de tais possibilidades.¹⁰

A noção que estou sugerindo está claramente relacionada à compreensão comum. As coisas no mundo normalmente nos afetam através da afetação da superfície do nosso corpo. Uma maneira óbvia de criar duas situações subjetivamente indistinguíveis para um sujeito é manter constante o estímulo próximo na superfície corporal, embora variando a origem causal do estímulo, e nesses casos, os fatos externos estão realmente fora do corpo. Assim, os casos da Terra Gêmea habitualmente discutidos tornar-se-ão casos da Terra Gêmea de acordo com a nova interpretação – como deveriam.

Além do mais, a minha proposta cobre os casos discutidos nas seções 2 e 3. No caso da meningite e em ambas as versões do caso da água, o ponto principal do argumento não é que os Gêmeos são idênticos molécula a molécula (não são); a estipulação crucial no cenário é que os Gêmeos estão em situações subjetivamente indistinguíveis.

Os críticos externistas da teoria cartesiana da mente geralmente identificam a descrição cartesiana das hipóteses do gênio maligno (ou dos cérebros numa cuba) como uma característica central da teoria. A ideia é que ainda que você fosse enganado pelo gênio maligno, ou fosse um cérebro numa cuba, seus pensamentos, não obstante, seriam os mesmos – apenas muitos deles seriam falsos. A viabilidade de toda hipótese depende do que identifiquei como a relação central na construção dos cenários da Terra Gêmea: a relação de indistinguibilidade subjetiva; pois se você fosse um cérebro numa cuba, *tudo pareceria o mesmo*. Os externistas se dividem sobre a questão do que dizer sobre

10 Poder-se-ia objetar que a tese da indistinguibilidade enquanto critério para a identidade dos estados mentais torna o internismo um fracasso por causa da intransitividade da indiscriminabilidade fenomênica. Não penso que esse ponto seja estabelecido: para um argumento a favor da transitividade dessa relação, veja Jackson e Pinkerton (1973) e Graff (2001). O principal problema que vejo com a negação da transitividade é, muito brevemente, este: a relação “mesma aparência” tem de ser transitiva (uma vez que se baseia na *identidade* da aparência). Negar a transitividade da indiscriminabilidade, portanto, compromete alguém com a negação da de que situações indiscrimináveis têm a mesma aparência. Isso é possível, mas penso que seja indesejável.

cérebros em cubas. Putnam concede a inteligibilidade da hipótese; o seu externismo é manifesto na afirmação de que contrário à suposição cartesiana, os pensamentos dos cérebros encubados seriam diferentes dos nossos pensamentos. Outros externistas questionam a inteligibilidade de todo o cenário. Em qualquer caso, o desacordo sobre as hipóteses do gênio maligno ou dos cérebros numa cuba entre internistas e externistas é um desacordo sobre o que implica a indistinguibilidade subjetiva, e isso torna imediatamente claro por que a teoria cartesiana é uma teoria internista.

Entender a relação entre os Gêmeos em termos de indistinguibilidade subjetiva é aplicável a outros tipos de argumentos externistas. O argumento de Putnam da divisão do trabalho lingüístico e o argumento de Burge a favor do externismo social envolvem imaginar duas comunidades lingüísticas onde o uso de certas expressões difere. Colocamos então um Gêmeo em cada uma dessas comunidades, e de acordo com o argumento, eles terão conceitos diferentes. Esses argumentos *não* seriam obviamente argumentos a favor do externismo se os Gêmeos de algum modo *registrassem* as diferenças relevantes no uso. A suposição crucial do cenário é novamente que se os Gêmeos fossem trocados contrafactualmente, *a situação seria indistinguível para o sujeito*.

5- Incompatibilidade e a compreensão comum

A significância dessa proposta está no fato de que ela sozinha ajuda a entender por que se deveria pensar que *ao contrário do internismo*, o externismo é incompatível com o auto-conhecimento ou o acesso privilegiado. Primeiro, vou argumentar que de acordo com a compreensão comum do externismo, podemos esperar nenhuma diferença entre o externismo e o internismo em relação ao autoconhecimento.

Suponha que aceitemos a compreensão comum do externismo que traça a linha divisória entre o interno e o externo na pele (ou no cérebro). A principal diferença então

entre o internista e o externista é sobre onde localizar os fatos aos quais o conteúdo de nossos estados mentais dependem:

Externismo:

Estar num estado mental com o conteúdo *C*
depende de/acarreta que
E (algum fato que está fora do corpo ou do cérebro do sujeito)

Internismo:

Estar num estado mental com conteúdo *C*
depende de/acarreta que
B (algum fato sobre o corpo ou sobre o cérebro)

A ideia de que o externismo é incompatível com o acesso privilegiado é geralmente articulada com a ajuda do contraste do nosso estatuto epistêmico no que diz respeito ao primeiro e segundo itens na tese externista. Assim, sabemos de algum modo especial (diretamente, ou *a priori*, ou com a autoridade da primeira pessoa, ou de algo parecido) que estamos num estado mental com conteúdo *C*, mas não sabemos dessa maneira especial que *E* é obtido. E como poderia algo que sabemos desse modo especial depender de ou acarretar algo que não sabemos desse modo especial? Os detalhes do argumento são preenchidos de acordo com aquilo que tomamos como “modo especial”, e de acordo com aquilo que tomamos com a natureza da “dependência” ou do “acarretamento” entre o primeiro e o segundo item. Atesta a formulação de Burge do problema em seu influente artigo defendendo a tese da compatibilidade:

O nosso problema é entender como podemos saber algum de nossos eventos mentais de maneira direta, não empírica, quando as identidades desses eventos dependem de sua relação com o ambiente. Uma pessoa não precisa investigar o ambiente para saber quais são seus pensamentos. Uma pessoa tem que investigar o ambiente para saber como é o seu ambiente. Isso não indica que os eventos mentais são aquilo que são independentemente do ambiente? (Burge, 1988, p. 650)

Mas se essa é de fato a fonte da preocupação com a compatibilidade, então o internista tem tanta razão para se preocupar quanto o externista. Considere a formulação do internismo acima: o mesmo contraste pode ser traçado entre o nosso estatuto epistêmico no que diz respeito ao primeiro e ao segundo item na tese. Certamente que não sabemos diretamente e não empiricamente os nossos estados cerebrais e, nem, sob uma descrição similar, os estados corporais pretendidos para individuar os nossos estados mentais. Descobrimos muitas coisas sobre o nosso corpo do mesmo modo que descobrimos coisas sobre o nosso ambiente: empiricamente e do ponto de vista da terceira pessoa – com a ajuda de raios-x, cirurgias, e amostra de tecidos. Se a *única e decisiva* diferença entre o internismo e o externismo é se eles alocam os fatos que individualizam o conteúdo mental dentro ou fora dos confins do corpo, não há razão para pensar que isso resultará em alguma diferença epistemológica entre as duas teses.

Burge faz uso desse *insight* em sua crítica à ideia de que o externismo é incompatível com o auto-conhecimento. De acordo com essa concepção, o argumento a favor da incompatibilidade tem o mesmo erro fundamental que o argumento de Descartes a favor da distinção real ente mente e corpo. Conceder que sabemos nossos estados mentais de um modo especial não implica que sabemos *tudo* sobre ele do mesmo modo especial – há ainda espaço para a ideia de que esses estados dependem de fatos sobre o corpo ou fatos sobre o ambiente.

Alguém pode saber quais são os seus eventos mentais e não saber, contudo, os fatos gerais relevantes sobre as condições para se individuar esses eventos. Simplesmente, não é verdade que o *cogito* nos dá o conhecimento das condições de individuação dos nossos pensamentos que nos permitem “destacar” as suas condições de individuação do ambiente físico. (Burge, 1988, p. 651)

Penso que Burge está essencialmente correto acerca disso.¹¹ É de algum modo intrigante, ainda mais porque muito tempo foi despendido na argumentação a favor e contra a incompatibilidade do auto-conhecimento e do externismo, se a solução do problema é tão simples. Sugiro a seguinte explicação: a solução é simples somente se nos apoiarmos na compreensão comum do externismo – nesse caso, os externistas e os internistas (materialistas) têm de fato tanta ou pouca razão para se preocupar com a compatibilidade e com o auto-conhecimento. Mas essa não é a última palavra no debate: pois de acordo com a minha nova compreensão, há uma diferença entre o internismo e o externismo em sua relação com o auto-conhecimento. Penso que uma confiança tácita em algo como a concepção que sugiro poderia explicar o sentimento persistente de que há um problema aqui. Antes de mostrá-la, deixe-me discutir brevemente outra versão popular do argumento da incompatibilidade.

6- O argumento de McKinsey

Uma forma frequentemente discutida do argumento da incompatibilidade segue um padrão primeiramente sugerido por Michael McKinsey. Eis o ponto principal do argumento:

[...] se você pudesse saber *a priori* que está num dado estado mental, e estar nesse estado acarretasse conceitual ou logicamente a existência de objetos externos, então você poderia saber *a priori* que o mundo externo existe. Uma vez que você obviamente não

11 Burge também apresenta uma teoria positiva do autoconhecimento que ele alega ser compatível com o externismo (para um desenvolvimento da teoria, veja também Burge, 1996). A essência da teoria é que pensamentos de segunda ordem como “Estou pensando agora que a água é úmida” são contextualmente auto-verificantes: uma vez que o conteúdo do pensamento de segunda ordem herda o conteúdo de primeira ordem “a água é úmida”, não há possibilidade de desacordo entre os dois conteúdos. Não penso que essa solução seja satisfatória. Não posso entrar em detalhes aqui, mas o principal problema parece ser que de acordo com a teoria de Burge, a correção dos pensamentos de segunda ordem é um resultado de seu caráter contextual; essencialmente do mesmo modo que não posso estar errado ao pensar que estou aqui. Esse último, porém, é compatível com eu não saber qualquer coisa sobre onde estou. É argumentável que no caso do autoconhecimento temos esse segundo tipo de conhecimento mais substancial – e, portanto, a teoria de Burge não dá conta de todo o escopo do autoconhecimento.

sabe *a priori* que o mundo externo existe, você também não pode saber *a priori* que está no estado mental em questão. É simples. (McKinsey, 1991, p. 16)

O argumento é uma *reductio*: a alegação de que o auto-conhecimento é *a priori*, combinada com a tese externista, conduz à conclusão inaceitável de que sabemos *a priori* certos fatos sobre o mundo externo. O que é essencial a esse argumento – mas, como veremos, também altamente controverso – é que a tese externista e sua aplicação específica no argumento deveria ser conhecida *a priori* num sentido suficientemente forte; de outro modo as pressuposições empíricas da tese podem explicar a natureza empírica da conclusão. Mas se isso pode ser mostrado ou não, o principal problema novamente é que se aceitarmos a compreensão comum, os internistas parecem ter tanta razão para se preocupar com o argumento de McKinsey quanto os externistas.

Para ver isso, apliquemos a *reductio* ao *internismo* como formulado acima (de acordo com a compreensão comum). De acordo com uma versão, temos a conclusão de que sabemos *a priori* que os nossos cérebros existem, o que é claramente falso. Dir-se-á que isso é porque o internismo no cérebro não é completamente *a priori*, sendo baseado na suposição empírica de que temos um cérebro. É bem verdade; o externismo fora do cérebro obviamente então também não é completamente *a priori*. Essa última teoria diz que os fatos fora de nosso cérebro individualizam os conteúdos mentais – e, portanto, a teoria assenta na suposição empírica de que temos um cérebro.

De acordo com a outra versão, temos a conclusão de que sabemos *a priori* que certo estado corporal existe. Isso levanta uma questão que não posso discutir aqui nos devidos detalhes: sabemos *a priori* que o nosso corpo existe? Se a resposta for não, então a conclusão desse argumento é novamente tão inaceitável quanto a conclusão tirada da tese externista, e novamente temos uma *reductio* contra a compatibilidade do internismo materialista e o acesso privilegiado. Agora, assim como antes, o conteúdo empírico da conclusão pode ser a consequência de alguma suposição empírica na tese internista. No entanto, a mesma suposição – nomeadamente a de que temos um corpo e isso marca

uma linha divisória relevante nos fatos localizados – também arruinará o caráter *a priori* da tese externista.

Mas talvez seja sugerido que a tese internista não incorrerá nas mesmas dificuldades que a posição externista, pois *sabemos a priori* que o nosso corpo existe. O que isso significa não é inteiramente claro: soa estranho dizer que sabemos que o nosso corpo existe independentemente de ou fora da nossa experiência. Portanto, precisaríamos de algum outro sentido de “*a priori*”, e é aqui que se poderia começar a duvidar de que nesse contexto seria útil apelar a essa noção. Em todo caso, parece provável que em qualquer interpretação do “*a priori*” que torne plausível que saibamos *a priori* que o nosso corpo existe, será defensável que sabemos *a priori* que objetos fora do nosso corpo existem. Por exemplo, pode-se sustentar que o fato da nossa existência incorporada é *parte do nosso esquema conceitual* e, por conseguinte, uma condição para qualquer experiência, mas não há algo de bizarro ou absurdo com a alegação de que o mesmo vale para a existência dos objetos materiais fora de nós.

Há muitas versões do argumento de McKinsey, e algumas delas podem evitar os problemas mencionados aqui, mas não posso fornecer uma discussão detalhada aqui. Deixem-me apenas dizer que não estou convencida de que o foco no suposto caráter *a priori* do autoconhecimento seja útil nesse contexto. Uma razão para isso é que a noção de *a priori* não desfruta de uma aceitação universal, uma vez que há diversos filósofos que foram convencidos por Quine de que não há conhecimento puramente *a priori*. No entanto, estivesse Quine correto sobre isso, parece-me que se poderia ainda argumentar a favor da natureza especial do autoconhecimento.¹² Segundo, ainda que resistamos à conclusão de Quine, não é sequer óbvio que o autoconhecimento seja *a priori* em qualquer compreensão plausível da noção. Tanto quanto posso ver, a melhor explicação do *a priori* já fornecida é em termos de analiticidade,¹³ mas não é aplicável ao autoconhecimento. Ao invés, somos deixados com termos de algum modo indefinidos como

12 Davidson parece ser um caso em questão: veja Davidson (1987).

13 Como por exemplo em Boghossian (1996).

“sem investigação empírica do mundo” ou “pelo pensamento apenas”. Na verdade, essas expressões dificilmente dizem mais do que há *alguma* diferença entre o modo que conhecemos o mundo e o modo que conhecemos nossos pensamentos.¹⁴

7- Externismo e Acesso Privilegiado

Se aceitarmos a minha compreensão do externismo, teremos um argumento que torna o acesso privilegiado incompatível com o externismo, mas não com o internismo. Primeiro, deveríamos obter clareza sobre as características do acesso privilegiado que gera a incompatibilidade. Já nos deparamos com a sugestão de que aquilo que torna o autoconhecimento privilegiado é o seu caráter *a priori*; mas pelas razões dadas acima, penso que não é útil. Há outras características do acesso privilegiado habitualmente defendidas, e a que recomendo nos focarmos é na *autoridade da primeira pessoa*. Ter autoridade de primeira pessoa sobre meus pensamentos não significa necessariamente infalibilidade sobre eles; significa apenas que estou numa melhor posição para saber meus próprios pensamentos do que outros.

O acesso privilegiado, quando caracterizado desse modo, é plausível primariamente para os pensamentos e as experiências ocorrentes. Explicar o conhecimento de nossas crenças, desejos ou intenções requer um relato mais complicado: fenômenos como o autoengano, a dificuldade de compreender ideias complexas ou os efeitos de envolvimento emocional forte sugerem que tais estados geralmente não são conhecidos com a autoridade de primeira pessoa. Por razões como essas, dificilmente alguém quereria sustentar que temos acesso privilegiado irrestrito a *todos* os nossos estados mentais. A característica marcante do externismo é a de que ele força uma limitação ao acesso privilegiado que é *fundamentalmente diferente* em caráter: surge no que diz respeito aos pensamentos e experiências ocorrentes mais simples, e

¹⁴ Outra fonte de problemas é o suposto caráter *a priori* da tese externista. Para dúvidas similares e um argumento convincente de que nenhuma noção do *a priori* servirá aos propósitos do argumento da incompatibilidade, veja Nuccetelli (1999).

não é explicável por esses fatos familiares da psicologia humana. Esse é um ponto importante que é geralmente negligenciado pelos externistas: eles simplesmente listam exemplos (como os acima) em que temos autoconhecimento limitado, e depois estendem facilmente essa limitação aos casos que são claramente bastante diferentes.

A incompatibilidade da autoridade de primeira pessoa e do externismo (na minha compreensão) é bastante clara. Os externistas sustentam que um sujeito em situações subjetivamente indistinguíveis poderia ter estados mentais diferentes. Mas a autoridade de primeira pessoa se estende apenas na medida em que as coisas são subjetivamente distinguíveis, isto é, distinguíveis do ponto de vista do sujeito. Se eu nunca tivesse notado a diferença entre essa situação e uma situação Gêmea, então certamente outras pessoas poderiam estar numa posição melhor para detectar a diferença. O internismo, por outro lado, não impõe restrições ao escopo da autoridade de primeira pessoa: pois de acordo com essa perspectiva, tudo que pudesse fazer diferença aos estados mentais de um sujeito deveria ser discriminável pelo próprio sujeito. O internista não tem de insistir que *todo* fato do qual os nossos pensamentos dependem – a existência do nosso cérebro, por exemplo – pode ser registrado pela autoridade de primeira pessoa. Não obstante, ele insiste que é legítimo afirmar que os fatos individualizam os conteúdos mentais apenas na medida em que *fazem diferença* ao modo como as coisas nos aparecem. Isso significa que qualquer diferença no conteúdo dos pensamentos deveria ser distinguível do ponto de vista do sujeito e, por conseguinte, permanecer no alcance do acesso privilegiado.

8- A Metafísica da Mente

Antecipo certa objeção. Alguns poderiam dizer que se defino externismo e acesso privilegiado *desse* modo, a incompatibilidade decorre; mas então o meu argumento é simplesmente uma petição de princípio. Penso que num sentido isso está correto: um modo de sumarizar a minha proposta é dizer que o externismo é uma tese sobre a natureza dos nossos acessos ou dos nossos pensamentos. Contudo, é importante ver que

nada do que eu disse põe fim ao debate internismo/externismo. Os argumentos apresentados neste artigo deixam diversas opções em aberto. Você pode escolher ignorar os pontos sobre o dualismo e do caso da meningite e sustentar a compreensão comum do externismo e internismo; mas nesse caso, você não deveria esperar estar numa posição especial sobre o autoconhecimento entre o externismo e o internismo. Alternativamente, você pode aceitar a minha concepção, insistir na verdade do externismo, e concluir que não temos autoridade de primeira pessoa sobre nossos pensamentos. Isso não significa necessariamente abandonar por completo a tese do acesso privilegiado; poder-se-ia ainda tentar-se dar conta da natureza privilegiada do autoconhecimento em termos de alguma outra característica. Finalmente, você poderia uma vez mais aceitar a minha concepção e argumentar que uma vez que a autoridade de primeira pessoa é uma característica essencial do conhecimento de nossos pensamentos, e que o externismo é incompatível com ela, o internismo vence.

Isso pode não satisfazer aqueles que vêem meu argumento como uma petição de princípio, de modo que vou concluir levando em conta uma objeção desse tipo que de fato vai ao cerne da questão. A objeção poderia correr do seguinte modo: “Um modo de explicar por que a questão do autoconhecimento e do externismo tem se mostrado tão difícil é apontar a natureza diversa das duas doutrinas. O externismo e o internismo, sendo teses sobre a individuação de conteúdo, são doutrinas *metafísicas* e, por conseguinte, deveriam ser postas em termos metafísicos, ao passo que a tese do acesso privilegiado é uma doutrina *epistemológica*, e deveria ser posta em termos epistêmicos. O que torna a demonstração da incompatibilidade difícil – ou mesmo impossível – é a dificuldade de extrair conseqüências epistêmicas de uma distinção metafísica. De fato, isso foi ilustrado muito bem pelo que foi dito sobre a dificuldade do argumento da incompatibilidade aceitar a compreensão comum. Não é surpreendente que essa *sua* demonstração da incompatibilidade seja tão fácil: pois você ilicitamente definiu o externismo e o internismo em termos de indistinguibilidade subjetiva, isto é, em *termos epistêmicos*. Mas assim você não traçou a distinção metafísica relevante”.

Manter separadas questões metafísicas e epistêmicas é geralmente um bom negócio, mas não penso que uma separação estrita seja praticável quando o nosso interesse é a *metafísica da mente*. O que é ter uma mente é inseparável do que é ter, por exemplo, experiências, e essa última é uma noção inteiramente epistêmica. O modo como a metafísica e a epistemologia se entrelaçam na prática filosófica sobre a mente pode ser ilustrado por inúmeros exemplos, de Descartes a Sellars. Isso é especialmente verdadeiro quando a questão é sobre conhecer a nossa própria mente; como disse Colin McGinn, “[...] não podemos *primeiro* formar uma concepção de mente e *depois* prosseguir na especificação dos modos pelos quais a mente é conhecida. Em suma, não há concepção epistemologicamente neutra da mente [...]” (McGinn, 1982, p. 7).

Em algumas discussões do externismo, é de fato claro que optar pelo externismo ou pelo internismo acaba na rejeição ou aceitação de alguma suposição epistêmica. Um bom exemplo é a discussão de John McDowell dos pensamentos dependentes de objetos (McDowell, 1986). Um pensamento é dependente de objeto se não pode existir sem que seu objeto exista. Na concepção original de Russell, a classe de tais pensamentos se limita aos pensamentos sobre os dados dos sentidos e aos pensamentos sobre nós próprios. McDowell sugere a expansão dessa classe a fim de incluir certos pensamentos sobre os objetos físicos externos e, por conseguinte, atingir uma concepção externista da mente. Ele deixa claro que

Com efeito, os resultados restritos de Russell decorrentes da recusa de se aceitar que possa haver uma ilusão na compreensão de uma frase aparentemente singular [...] (138)

Se suspendermos a restrição de Russell, abrimos a possibilidade de que um sujeito possa estar em erro sobre o conteúdo de sua própria mente [...] (145)

O que motiva o externismo de McDowell é a convicção de que ele fornece o único modo de dar conta da relação entre a mente e o mundo: que os nossos pensamentos são *acerca de* objetos no mundo externo. Mas ele também percebe que adotar o externismo envolve imediatamente uma restrição ao acesso aos nossos pensamentos. E penso que

essa oposição ao externismo pode muito bem emergir da rejeição do compromisso com essa questão.¹⁵

Referências

- Blackburn, S. (1984): *Spreading the Word*, Oxford: Clarendon Press.
- Boghossian, P.A. (1996): “Analyticity Reconsidered”, *Noûs* 30/3, 360–391.
- Boghossian, P.A. (1997): “What the Externalist Can Know A Priori”, *Proceedings of the Aristotelian Society* XCVII/2, 161–175 (também em Wright-Smith-MacDonald, 1998).
- Burge, T. (1982): ‘Other Bodies’, in Andrew Woodfield (ed.), *Thought and Object* (pp. 97–120), Oxford: Clarendon Press.
- Burge, T. (1986): “Cartesian Error and the Objectivity of Perception”, in Pettit-McDowell 1986, 117–136.
- Burge, T. (1988): “Individualism and Self-Knowledge”, *Journal of Philosophy* 85, 649–663.
- Burge, T. (1996): “Our Entitlement to Self-Knowledge”, *Proceedings of the Aristotelian Society* XCVI, 91–116.
- Davidson, D. (1987): “Knowing One’s Own Mind”, *The Proceedings and Addresses of the American Philosophical Association* 60, 441–458, reimpresso em Cassam, Quassim (ed.) 1994, *Self-Knowledge* (pp. 43–64), Oxford University Press.
- Davies, M. (1998): “Externalism, Architecturalism and Epistemic Warrant” in Wright-Smith-MacDonald 1998, 321–361.
- Graf, D. (2001): “Phenomenal Continua and the Sorites”, *Mind* 110, 905–935.
- Jackson, F. e Pinkerton, R.J. (1973): “On an Argument against Sensory Items”, *Mind* 82, 269–272.

¹⁵ O trabalho neste artigo foi apoiado pelo Philosophy of Language Research Group da Hungarian Academy of Sciences. Gostaria também de agradecer a Tim Crane e Barry C. Smith pelos comentários.

- Jackson, F. e Pettit, P. (1996): “Functionalism and Broad Content” em A. Pessin and S. Goldberg (eds.), *The Twin Earth Chronicles* (pp. 219–230), M. E. Sharpe.
- McCulloch, G. (1995): *The Mind and its World*, London e New York: Routledge.
- MacDonald, C. (1998): “Externalism and Authoritative Self-Knowledge” in Wright-Smith-MacDonald 1998, 124–154.
- McDowell, J. (1982): “Criteria, Defeasibility and Knowledge”, em Dancy Jonathan (ed.) 1988, *Perception Knowledge* (pp. 209–219), Oxford University Press.
- McDowell, J. (1986): “Singular Thought and the Extent of Inner Space”, em Pettit-McDowell 1986, 136–168.
- McGinn, C. (1982): *The Character of Mind*, Oxford, New York: Oxford University Press.
- McKinsey, M. (1991): “Anti-Individualism and Privileged Access”, *Analysis* 51(1), 9–16.
- McLaughlin, B.P. e Tye, M. (1998): “Externalism, Twin-Earth and Self-Knowledge”, em Wright-Smith-MacDonald 1998, 285–320.
- Nuccetelli, S. (1999): “What Anti-Individualists Cannot Know A Priori”, *Analysis* 59(1), 48–51.
- Pettit, P. e McDowell, J. (1986): *Subject, Thought and Context*, Oxford: Clarendon Press.
- Pettit, P. (1986): “Broad-Minded Explanation and Psychology” em Pettit-McDowell 1986, 17–58.
- Putnam, H. (1975): “The Meaning of ‘Meaning’”, em *Mind, Language and Reality*, Cambridge University Press.
- Segal, Gabriel (2000): *A Slim Book about Narrow Content*, Cambridge, MA: The MIT Press.
- Wright, C., Smith, B.C. e MacDonald, C. (1998) (eds.): *Knowing Our Own Minds*, Oxford University Press.

Como explicar o fenómeno da vagueza?

Domingos Faria

Universidade de Lisboa e Lancog

Resumo: O meu objetivo com este artigo é fazer uma introdução a um problema de lógica filosófica: a vagueza. Para isso, exponho o paradoxo sorites, assim como explico uma teoria que tenta resolver este problema – a teoria epistémica – e, por fim, examino uma objeção a essa teoria.

Palavras-chave: Vagueza, Sorites, Teoria Epistémica.

Abstract: My aim in this paper is to do an introduction to a problem of philosophical logic: vagueness. For that, I will present the sorites paradox, as well as I will explain a theory that it tries to solve this problem – the epistemic theory – and, lastly, I will survey an objection to that theory.

Key-words: Vagueness, Sorites, Epistemic Theory.

Numa loja de roupa uma funcionária questionou-me se precisava de ajuda para comprar alguma coisa. Apontei para uma camisola e perguntei: “Acha que aquela camisola vermelha me fica bem?”. A funcionária respondeu que a camisola não era claramente vermelha, apesar de também não ser claramente cor-de-laranja. Olhando melhor parece que ela tinha razão. Então, afinal a camisola que pretendo comprar é vermelha ou não?

O predicado “vermelho” parece demasiado vago para que “aquela camisola é vermelha” seja reconhecida como verdadeira ou falsa. E o predicado “vermelho” é vago, assim como “monte”, “alto” ou “magro”, uma vez que admite casos de fronteira, ou seja, casos onde não é claro ou em que não sabemos se o predicado se aplica ou não, mesmo estando em condições ótimas de observação ou tendo as informações relevantes que normalmente são suficientes para atribuir um determinado predicado. Deste modo,

posso saber exatamente quantos grãos de areia estão amontoados e mesmo assim não ser capaz de decidir se tenho ou não um monte de areia. Ou posso medir uma determinada pessoa e não conseguir determinar se ela é alta ou não. Ou ainda posso estar em condições ótimas de observação e não saber se aquela camisola que pretendo comprar é vermelha ou cor-de-laranja. Deste modo, parece que a extensão destes termos vagos não está claramente determinada, admitindo zonas de penumbra, tendo fronteiras turvas ou imprecisas.

Ao falar de predicados vagos não estou a referir-me a predicados ambíguos. Um predicado é ambíguo quando se aplica a várias coisas, é polissémico, como “banco” que se pode referir a uma instituição financeira ou a um objeto em que me posso sentar. No entanto, posso eliminar aquilo que é ambíguo e continuar a ter predicados vagos. E se repararmos bem a linguagem que utilizamos quotidianamente está repleta de palavras vagas. Assim, surge a questão: como se poderá explicar este fenómeno da vagueza? Será a vagueza uma propriedade da linguagem, ou do mundo, ou antes de nós próprios?

Os problemas adensam-se pelo facto de a vagueza, ao admitir zonas de penumbra, parecer colocar em causa o princípio de bivalência ou o terceiro excluído, o que obrigaria a abandonar ou pelo menos a rever a lógica clássica. Além disso, é por causa dos predicados vagos que surgem os paradoxos de sorites, tal como o seguinte.

No mesmo local onde pretendia comprar a camisola encontrava-se um pintor que se preparava para alterar a decoração da parede da loja. Para isso ele trouxe consigo duas latas de tinta do mesmo tamanho: uma vermelha e outra amarela. Não pretendia ter a parede toda vermelha, nem toda amarela; pelo contrário, o seu objetivo era começar por traçar a parede com um tom claramente vermelho e gradualmente chegar a um tom claramente cor-de-laranja. Mas como fazer isso? Começa por colocar uma gota de tinta amarela bem misturada na lata de tinta vermelha e com isso continuou a ter a tinta vermelha. Então quando é que temos tinta cor-de-laranja? Certamente duas gotas de tinta amarela misturadas em tinta vermelha não fazem tinta cor-de-laranja. E três gotas de tinta amarela? Também não parecem fazer a tinta cor-de-laranja. E quatro? E cinco? Então quando deixamos de ter tinta vermelha e passamos a ter tinta cor-de-laranja? Este seu raciocínio gerou um paradoxo que pode ser formalmente representado do seguinte modo:

- (1) Uma gota de tinta amarela bem misturada numa lata de tinta vermelha não faz tinta cor-de-laranja.
- (2) Se não se pode fazer tinta cor-de-laranja com n gotas de tinta amarelas bem misturadas numa lata de tinta vermelha, então não se pode fazer tinta cor-de-laranja com $n+1$ gotas de tinta amarelas bem misturadas numa lata de tinta vermelha.
- (3) Logo, continuamos a não fazer tinta cor-de-laranja mesmo quando misturamos todas as gotas da lata de tinta amarela com a lata de tinta vermelha.

Será que o pintor pode aceitar a conclusão? O raciocínio é válido e as premissas parecem verdadeiras. A premissa (1) é incontroversa. A premissa (2) também parece plausível, pois se uma gota não faz a diferença então duas gotas também não fazem a diferença, e se duas gotas não fazem a diferença então igualmente três gotas não fazem a diferença... Portanto, esta premissa, que obedece ao princípio de tolerância, realça que uma gota amarela a mais da anterior parece não fazer a diferença na tinta deixar de ser vermelha. Porém, a conclusão é falsa, uma vez que se misturarmos bem toda a tinta da lata amarela com toda a tinta da lata vermelha obtemos tinta cor-de-laranja. Como resolver este paradoxo? O que haverá de errado neste argumento?

Uma forma de resolver o paradoxo, que fornece do mesmo modo uma explicação para o fenómeno da vagueza, é recorrer à teoria epistémica de Williamson (1992: 145-162). Para este filósofo a vagueza não é uma propriedade da linguagem ou da semântica (em que é o resultado de um significado incompleto ou deficiente – como no supervalorativismo – sendo os predicados vagos indeterminados), nem é uma propriedade do mundo (em que as coisas por si próprias – *de re* – seriam inerentemente vagas), mas sim é uma propriedade de nós mesmos (sobretudo da nossa ignorância e limitações epistémicas). Deste modo, para Williamson, a “vagueza é um tipo de ignorância” (1992: 145). Isto porque existem fronteiras claras que delimitam o domínio de aplicação dos predicados vagos, tal como no caso dos predicados concretos; mas nos predicados vagos o nosso mecanismo cognitivo é insuficiente para conhecermos onde está a fronteira. Ou seja, somos ignorantes sobre onde reside a fronteira.

Atendendo a esta teoria, pode-se considerar que a premissa (2) do sorites é falsa. Pois, se a conclusão é falsa e o argumento é válido, então tem de haver algures uma fronteira nítida que delimita o número de gotas de tinta amarela que continuam a fazer tinta vermelha daquele número de gotas que não fazem tinta vermelha. Ou seja, haverá um conjunto de n gotas que fazem tinta vermelha, mas haverá um conjunto de $n+1$ gotas que não fazem tinta vermelha. Assim, não se gera o paradoxo se rejeitarmos o princípio de tolerância e adotarmos o princípio de margens de erro (Williamson 1992: 161). Adotando as margens de erro para o predicado de cor vago ficaríamos com algo como: Se X e Y diferem em tons de cor por menos do que uma pequena constante e X é conhecido como vermelho, então Y é vermelho. Ora, quando chegarmos a uma dada altura em que não sabemos se um determinado tom de cor com n gotas de tinta é vermelho ou não, então não podemos afirmar que o próximo tom com $n+1$ gotas de tinta será vermelho – estamos numa situação de ignorância. Em suma, existe uma fronteira que delimita o vermelho do não vermelho, mas não sabemos onde reside essa fronteira nítida.

O mesmo acontece quando o pintor começa por fazer um fino traço vertical na parede da loja com tinta da lata vermelha misturada com uma gota amarela, depois outro traço com duas gotas, seguido por um novo traço com três gotas, e assim sucessivamente até termos claramente traços cor-de-laranja. As paredes da loja ficam completamente pintadas e podemos presenciar a um espectro de cor que vai gradualmente de vermelho a cor-de-laranja. De acordo com a teoria epistémica da vagueza há uma fronteira clara neste espectro que delimita os traços vermelhos dos traços cor-de-laranja; contudo, por causa do nosso mecanismo cognitivo e devido aos limites dos nossos sentidos para distinguir cores bastante aproximadas, somos ignorantes em detetar essa fronteira.

É importante salientar que se eu aponto para algo que seja caso de fronteira e expresso que “aquela camisola é vermelha” ou “aquela região da parede é vermelha” não estou a proferir algo indeterminado, mas sim algo que é ou verdadeiro ou falso; porém, devido à constituições cognitivas do ser humano, não estou em circunstâncias de saber se aquelas expressões são verdadeiras ou falsas. Com isto, e ao contrários das alternativas teóricas, não é preciso abandonar a lógica clássica nem a reformular.

Mas, será esta uma teoria plausível? Uma das formas de testar a sua solidez é apresentar objeções (Sainsbury 2009: 51) como a seguinte:

- (4) Se um predicado se aplica a uma propriedade manifesta (uma propriedade que pode estar claramente presente em certas condições), então em condições ótimas de manifestação, se há um facto sobre se essa propriedade está presente, esse facto é ele mesmo detetável.
- (5) Podemos ver uma região de uma parede colorida (uma fronteira de “vermelho”) em condições de observação ótimas, sem sermos capazes de detetar a presença ou ausência de vermelho.
- (6) Logo, não há nenhum facto em que essa região da parede ou é vermelha ou não o é.

Será este um bom argumento? Como o podemos avaliar? O argumento parece defender que não há uma fronteira nítida entre vermelho e não vermelho, de modo que não podemos atribuir a um caso de fronteira de “vermelho” um facto como sendo verdadeiro ou falso, constituindo antes algo indeterminado, isto é, nem verdadeiro nem falso. Assim, o termo vago “vermelho” numa zona de penumbra falha no valor de verdade bivalente. E isto porque se existe um predicado que se aplica a uma propriedade e a um facto manifestos, então estes são detetáveis em condições ótimas de manifestação. Mas, ao olharmos em condições ótimas para uma parte da parede da loja, para um caso de fronteira de “vermelho”, não somos capazes de detetar se a parede é vermelha ou não. Portanto, não é o caso que tal parte da parede seja ou vermelha ou não.

Não me parece que este argumento contra a teoria epistémica seja sólido, pois penso que parte de pelo menos uma premissa que é falsa e também conduz a uma conclusão falsa. Pode-se negar a premissa (4) afirmando que um predicado aplica-se a uma propriedade manifesta, mas ainda assim não é detetável mesmo em condições ótimas de manifestação. E isto parece-me plausível se imaginar que vou assistir ao próximo clássico de futebol entre Portugal-Brasil e, no minuto anterior ao início do jogo, olho para as bancadas e transmito ao meu colega do lado que:

(*) O número de pessoas a assistir ao jogo é ímpar.

No entanto, mesmo em condições ótimas de observação, (*) não é detetável por causa do nosso limite cognitivo em contar um grande número de pessoas. Deste modo, existe um facto claramente presente (o número de pessoas ou é par ou é ímpar), mas não somos capazes de detetar tal facto (o mesmo se passa em relação ao “vermelho”). Além disso, vamos supor que o número de pessoas a assistir ao jogo é 49999; assim, a minha afirmação (*) seria verdadeira. Porém, não posso dizer que conheci (*), pois não seria certa fiavelmente, mas apenas por sorte.

Do mesmo modo, penso que não podemos apropriadamente ir da premissa (5) para a conclusão (6). Pois, do facto de não conseguirmos detetar a presença ou ausência de vermelho, mesmo em ótimas condições, não se segue que não exista ou vermelho ou não vermelho, tal como do facto de não conseguirmos detetar se o número de pessoas no estádio é par ou ímpar não se segue que não exista um número de pessoas no estádio que é par ou ímpar. Aliás, podíamos utilizar a mesma estrutura argumentativa da objecção para provar absurdamente que não é o caso que o número de pessoas que assistem ao jogo é par ou é ímpar:

- (7) Se um predicado se aplica a uma propriedade e facto manifesto, então estes são detetáveis em condições ótimas de manifestação.
- (8) Podemos olhar em condições ótimas para as bancadas de um estádio sem sermos capazes de detetar se o número de pessoas que assistem ao jogo é par ou ímpar.
- (9) Logo, não há nenhum facto em que o número de pessoas que assistem ao jogo ou é par ou não o é.

Mas, como é óbvio que o número de pessoas tem de ser par ou ímpar, então esta estrutura argumentativa é errada e alguma premissa tem de ser falsa (nomeadamente a primeira). Igualmente as conclusões (6) e (9) parecem absurdas, uma vez que estão a negar o princípio de bivalência originando, deste modo, uma contradição. Por outras palavras (6) afirma que “não é o caso que aquela região da parede é vermelha ou aquela região da parede não é vermelha”. Ora, aplicando uma das leis de Morgan teríamos a

afirmação de que “aquela região da parede não é vermelha e aquela região da parede é vermelha”, o que constitui uma contradição sendo, assim, absurdo negar o princípio de bivalência (Williamson 1992: 145-146).

Portanto, penso que o argumento não consegue mostrar que não há nenhum facto claramente presente em que aquela parte da parede da loja ou é vermelha ou não o é. Aliás, tal como podemos arranjar um mecanismo para contar o número de pessoas que entram no estádio de futebol determinando com segurança se é par ou ímpar, também seria logicamente possível pensar numa máquina que analise com rigor os comprimentos de onda de cor e determine fiavelmente se estamos perante uma cor vermelha ou não. Agora pelo facto de nós próprios não conseguirmos pela observação detetar essas fronteiras não quer dizer que tais fronteiras não existam – apenas somos ignorantes sobre onde elas residem. Assim, considero que esta não é uma boa objecção contra a teoria epistémica.

Por esta objecção falhar não se segue que a teoria epistémica é plausível; pois, parece contraintuitivo, por exemplo, admitir que existe uma fronteira clara entre criança e adolescente, tendo em conta o que estes predicados significam. Haverá por ventura um único segundo de idade que delimita a passagem entre ser criança e ser adolescente?! Não haverá neste caso uma fronteira indeterminada? Desta forma, fico com dúvidas se esta é a melhor teoria para explicar o fenómeno da vagueza.

Bibliografia

- Keefe, Rosanna (2000) “The epistemic view of vagueness”. In *Theories of Vagueness*. Cambridge: Cambridge University Press.
- Ruffino, Marco (2003) “Vagueza de re”. In *O Que Nos Faz Pensar*, nº 17.
- Sainsbury (2009) “Vagueness: the paradox of the heap” In *Paradoxes*. New York: Cambridge University Press.
- Santos, Pedro (2006) “Sorites”, in *Enciclopédia de Termos Lógico-Filosóficos*. São Paulo: Martins Fontes.
- Santos, Ricardo (2010) “O Supervalorativismo e a Vagueza de Ordem Superior”. In *Paradoxos: Dedutivos, Indutivos e Práticos*, <http://www.filosofia.uevora.pt/paradoxos/trabalhos.htm>

- Williamson, Timothy (1992) “Vagueness and Ignorance”. In *Proceedings of the Aristotelian Society, Supplementary Volumes*, Vol. 66, pp.145-162.
- Williamson, Timothy (1994) “Vagueness as ignorance”, “Inexact knowledge”. In *Vagueness*. London and New York: Routledge.

Resenha de McGinn, Marie (2009). *Elucidating the Tractatus – Wittgenstein’s Early Philosophy of Logic and Language*. Oxford University Press: New York.

Fernando Spósito Yokoyama*

UNICAMP

Em seu livro *Elucidating the Tractatus*¹, Marie McGinn tem como principal objetivo apresentar e defender uma interpretação das teses acerca da relação entre linguagem e realidade desenvolvidas no *Tractatus Logico-Philosophicus* de Ludwig Wittgenstein. Ela denomina sua interpretação de “anti-metafísica” e a situa entre duas leituras extremas e opostas presentes na literatura secundária. Uma delas, que a autora chama de *realista*, é defendida por David Pears, Max Black e G.E.M. Anscombe, entre outros, e constitui a leitura mais tradicional da obra. A segunda foi desenvolvida por intérpretes como Cora Diamond e James Conant, que rejeitaram a interpretação realista inaugurando a que ficou assim conhecida como a leitura *resoluta* do *Tractatus*. McGinn associa a leitura anti-metafísica que ela defende a intérpretes como Hidé Ishiguro, Brian McGuinness, Rush Rhees, e Peter Winch, não pleiteando para si originalidade em muitas das ideias que apresenta, embora desenvolva essa abordagem interpretativa de uma maneira mais detalhada do que fizeram esses outros autores, sendo esse, portanto, um dos principais méritos de seu livro.

McGinn trata também de alguns temas que não se relacionam tão diretamente à polêmica mencionada acima. Por exemplo, ela traça um panorama do contexto filosófico no qual Wittgenstein estava situado quando iniciou suas investigações que culminaram no *Tractatus*. Deste modo, a autora procura explicitar as principais dificuldades que ele identificou nas obras de Frege e Russell, indicando assim os problemas que o motivaram em suas investigações iniciais. Outros temas que não dizem respeito propriamente ao que há de distintivo

*O presente trabalho foi realizado com apoio do CNPq, Conselho Nacional de Desenvolvimento Científico e Tecnológico – Brasil.

¹ McGINN, Marie - *Elucidating the Tractatus: Wittgenstein’s Early Philosophy of Logic and Language*. Oxford: Clarendon Press, 2006.

em sua interpretação anti-metafísica da obra compreendem o tratamento que o *Tractatus* dá às proposições moleculares e à natureza da inferência lógica. Uma vez que sua abordagem desses assuntos não difere de uma maneira decisiva do que já foi exaustivamente trabalhado por outros intérpretes, não me ocuparei aqui com o que McGinn tem a dizer sobre eles. Meu foco está dirigido, sobretudo, aos aspectos essenciais da leitura anti-metafísica, tema que ocupa a autora na maior parte do seu livro. Ao final, pretendo apontar os méritos dessa sua interpretação bem como examinar alguns aspectos dessa leitura que, penso eu, são problemáticos ou não receberam atenção suficiente.

Como foi dito acima, McGinn situa sua leitura entre duas linhas interpretativas extremas e opostas: as leituras realista e resoluto. Os partidários da leitura realista sustentam que, segundo o *Tractatus*, nossa capacidade de representar a realidade por meio da linguagem estaria fundada em relações diretas entre nomes e objetos, que são, ambos, elementos simples e indefiníveis de seus respectivos domínios. O que caracteriza essa interpretação é a ideia de que tais objetos devem ser entendidos como constituindo um domínio anterior e independente da linguagem. Eles afirmam, assim, que, a partir de uma investigação lógica sobre a natureza essencial da linguagem, Wittgenstein teria deduzido teses ontológicas acerca de aspectos essenciais da realidade, concebida como algo independente de nossos meios de representá-la.

Essa linha interpretativa foi rejeitada por autores como Cora Diamond e James Conant. Estes últimos dão ênfase ao caráter anti-teórico e ao status de contra-senso (*unsinnig*) que o próprio Wittgenstein atribuiu à sua filosofia e aos aforismos do *Tractatus*, respectivamente. De acordo com a leitura *resoluta* que eles sustentam, devemos entender as passagens em que Wittgenstein parece estar enunciando teses metafísicas acerca da relação entre linguagem e realidade como parte de uma tarefa exclusivamente *terapêutica*: quando, no final do livro, ele reconhece seus aforismos como contra-sensos, entendemos que seu verdadeiro e único objetivo foi mostrar a impossibilidade de qualquer tentativa de enunciar teses acerca do funcionamento da linguagem, bem como de sua relação com a realidade. Seríamos então deixados sem qualquer conhecimento positivo acerca da

natureza da linguagem ou realidade. Desse modo, esses intérpretes procuram fazer justiça ao caráter anti-teórico que Wittgenstein atribui à sua filosofia, como quando afirma, em 4.112, que filosofia não é uma teoria e que seu resultado não são “proposições filosóficas”. A interpretação que McGinn desenvolve no seu livro aparece como uma alternativa a essas duas leituras extremas, que ela considera insatisfatórias.

Ao situar sua leitura como ocupando uma posição intermediária entre essas duas interpretações extremas, McGinn procura combinar certas ideias de cada uma, que ela julga adequadas. Assim como os partidários da leitura *resoluta*, ela procura desenvolver uma interpretação compatível com as passagens em que Wittgenstein afirma que a filosofia não consiste em um corpo teórico de teses substanciais. Ela observa a seguir (p.13) que, já nas ‘Notes on logic’ (texto anterior ao *Tractatus*) Wittgenstein afirmara que a filosofia é puramente *descritiva*. Segundo McGinn, o que ele estaria excluindo é a presença, na filosofia, de hipóteses acerca da natureza da relação entre linguagem e mundo, concebidos como domínios independentes entre si. Os resultados da filosofia devem ser alcançados por uma investigação *interna* à linguagem, uma vez que é apenas o *uso* das expressões em proposições com sentido que pode revelar a essência do sentido proposicional. Para a autora, esse seria o teor da afirmação de Wittgenstein, em 4.112, de que uma obra filosófica consiste essencialmente de “elucidações”: “The force of the claim that philosophy is ‘purely descriptive’ is to be understood, in part at least, as a call to allow *language itself* to reveal how it functions” (p. 20; meu grifo). A principal consequência que ela extrai disso é a recusa da ideia, partilhada pelos intérpretes realistas, de que as investigações lógicas do *Tractatus* tenham levado à inferência de teses ontológicas acerca de uma realidade independente.

Porém, ao contrário do que propõe a leitura *resoluta*, McGinn afirma que a tarefa puramente elucidativa e descritiva da filosofia, da maneira como é executada no *Tractatus*, leva a descobertas *positivas* acerca do funcionamento da linguagem. Para ela, assim como para os intérpretes realistas, os aforismos da obra não devem ser tomados como absurdos vazios a serem completamente descartados no final do livro. Antes, eles transmitem um conhecimento genuíno acerca de como a

linguagem simboliza, ainda que esse conhecimento tenha sido obtido através de um exame exclusivamente interno à ela.

Por sustentar que o *Tractatus* nos oferece, afinal, um conhecimento positivo acerca do funcionamento da linguagem, McGinn incorpora à sua interpretação alguns aspectos importantes que também estão presentes na leitura realista que ela ataca. Por exemplo, ela admite que, segundo o *Tractatus*, se uma proposição qualquer é capaz de expressar um sentido, então sua análise completa deve necessariamente poder atingir um nível formado por proposições elementares, compostas de nomes simples e indefiníveis (p. 109-113). Para ela, assim como na leitura realista, a razão disso decorre da demanda tractariana de que o sentido de uma proposição não pode depender de nada contingente, ou seja, do valor de verdade de outra proposição. Wittgenstein teria postulado, então, que deve haver nomes que significam objetos simples e que, portanto, nunca podem falhar em ter significado, de modo que o sentido das proposições em que eles ocorrem passa a não depender mais de algo contingente. O que há de distintivo na interpretação de McGinn é que, segundo ela, essa é uma investigação que jamais ultrapassa os limites da linguagem. As teses apresentadas no *Tractatus* diriam respeito exclusivamente ao que está pressuposto em nossa posse de um sistema de representação por meio do qual construímos proposições com sentido. McGinn diz:

Wittgenstein is not attempting to deduce something about reality that lies outside language, but is attempting to clarify the logical order that is essential to any system in which possible states of affairs are represented (p. 100).

Enquanto para a leitura realista, a necessidade de nomes indefiníveis teria levado Wittgenstein a incorporar ao seu sistema teses ontológicas que dizem respeito a uma realidade anterior e independente da linguagem, McGinn, ao contrário, afirma: “We are not being directed to an order that is there independently of language, but to the order that is essential to *the depiction of states of affairs*” (p. 100; meu grifo).

Esse modo de conceber a natureza das teses tractarianas leva a um entendimento particular da teoria afigurativa do sentido da proposição,

apresentada no *Tractatus*, que afirma que uma proposição pode expressar um sentido por ser uma *figuração* do estado de coisas que ela enuncia. Segundo uma leitura realista, o contato entre uma proposição e o estado de coisas que ela afigura se realiza por meio de ligações *diretas* e *externas* entre nomes e objetos. Seria por meio dessas ligações que a proposição se conectaria ao estado de coisas afigurado, tomado como algo externo e absolutamente independente da linguagem. Por outro lado, para McGinn, uma consequência de conceber as investigações tractarianas como restritas a um domínio *interno* a linguagem é exatamente a recusa de que se possa postular ligações externas desse tipo: “There is, on the interpretation to be developed here, no attempt in the *Tractatus* to get outside language and explain how it hooks up with the world” (p. 79).

Com isso, McGinn pretende fazer justiça ao chamado “princípio do contexto”, segundo o qual um nome só tem significado no contexto de uma proposição, ao qual Wittgenstein adere explicitamente no aforismo 3.3. Para McGinn, ao afirmar isso Wittgenstein estaria recusando que o significado de um nome possa ser explicado por ligações diretas com objetos que se dariam independentemente do *uso* desse nome em contextos proposicionais. O que confere significado aos nomes e, por conseguinte, sentido à proposição, é o *uso* deste nome em figurações que podem ser comparadas com a realidade a fim de se decidir sua verdade ou falsidade. Isso, insiste McGinn, é algo que se dá *internamente* à linguagem. O emprego de uma determinada expressão pressupõe o domínio das “regras de projeção” (pp. 79-85), cuja aplicação a sinais proposicionais (ex: sons e traços no papel) mostra *qual* estado de coisas a proposição está afigurando. Tais regras dizem respeito apenas ao sistema de representação enquanto tal: não estão fundadas em um domínio independente de objetos com os quais a linguagem manteria uma relação externa. Antes, é a identidade do objeto ao qual um nome se refere que depende totalmente do *uso* desse nome segundo regras de projeção que são explicitadas através de uma investigação exclusivamente interna à linguagem. Diz McGinn:

The object for which a name stands is not something that exists over and against language, but is what we grasp when we grasp the meaning of a name, that is when we grasp the contribution that a name makes to determining the sense of a class of propositions (p. 114).

Esse aspecto de sua leitura é reafirmado quando ela diz que é o *sentido* da proposição que a conecta com a realidade, e não conexões individuais entre nomes e objetos: “It is the possibility of comparing a picture with reality for correctness or incorrectness that constitutes a picture’s connection with reality. It is, in other words, the sense of the picture that connects it with reality (...)” (p. 89). A conexão da proposição com a realidade decorre do fato de que seu valor de verdade depende de que algo seja ou não o caso no mundo. Vemos assim que, apesar de McGinn recusar que os objetos do *Tractatus* constituam um domínio independente, ela não nega que, segundo o *Tractatus*, deva haver um tal domínio independente da linguagem, mas apenas na medida em que esse domínio é capaz de determinar o valor de verdade das proposições: “The subject [isto é, o usuário da linguagem] does not create the world: which states of affairs exist or do not exist is independent of the subject, and established only a posteriori” (p. 273).

Ao admitir, porém, uma realidade independente que verifica ou falsifica proposições, McGinn não pretende comprometer sua leitura com teses ontológicas. Estas últimas surgem quando se afirma, como em uma leitura realista, que esse domínio independente possui a mesma “forma lógica” que as proposições que ele verifica ou falsifica, ou, mais ainda, quando se afirma que tal domínio *impõe* essa sua forma à linguagem. Nesse sentido, todas as propriedades essenciais da linguagem enunciadas no *Tractatus* pertenceriam à linguagem apenas na medida em que esta, para poder representar, “duplica” as propriedades essenciais da realidade representada. Ao afirmar que as investigações do *Tractatus* dizem respeito apenas ao que é essencial ao nosso sistema de representação enquanto tal, McGinn pretende recusar esse modelo realista. Para ela, não é a realidade, concebida como algo independente, que impõe sua forma à linguagem, mas, antes, é a estrutura essencial da linguagem que determina a forma do mundo nela

representado: “We can think of this logical order as what language and reality essentially have in common insofar as the logical order of the method of projection is essentially *preserved* in what is projected” (p. 100; meu grifo). Deste modo, McGinn faz uma releitura das passagens iniciais da obra, que, aparentemente, estariam comunicando teses metafísicas acerca de uma ordem a priori que a realidade possuiria independentemente de nossos meios de representá-la. Para McGinn, o que Wittgenstein está fazendo é simplesmente traçar a ordem lógica essencial de um sistema de representação que expressa proposições verdadeiras ou falsas: “What we come to see is that what Wittgenstein is doing in these remarks is not metaphysics but logic” (p. 137).

Tendo apresentado as ideias essenciais da leitura anti-metafísica que McGinn defende em seu livro, passo agora a fazer agora algumas observações acerca dos aspectos que considero positivos e também do que considero problemático em sua interpretação. A meu ver, McGinn tem êxito em sua tarefa de desenvolver essa linha interpretativa de uma maneira mais completa e detalhada do que outros de seus proponentes, ao procurar extrair todas as implicações dessa interpretação, e encontrar, de maneira convincente, suporte no texto de Wittgenstein. Passagens como 3.3-3.318 e 3.326-3.331 dão apoio textual à ideia central de sua leitura: a de que significado, no *Tractatus*, diz respeito ao *uso* de um nome em proposições com sentido, e não depende de supostas ligações externas com objetos, que ocorreriam fora de contextos proposicionais.

McGinn faz também uma interessante conexão entre esses aspectos de sua leitura e as passagens do *Tractatus* que tratam do solipsismo, nas quais Wittgenstein afirma que a verdade do solipsismo consiste no fato de que: “*Os limites de minha linguagem* significam os limites do meu mundo” (5.6). Segundo McGinn, isso confirmaria sua ideia central de que, no *Tractatus*, o mundo possui a forma que possui apenas na medida em que é representado na linguagem: “(...) the point of TLP 5.6 is that the world is simply what is described in the true propositions of my language, ‘that language which alone I understand” (p. 272). Passagens como 5.6, mais toda a discussão acerca do solipsismo, podem ser

tomadas como favoráveis à interpretação anti-metafísica, uma vez que parecem ir contra a ideia de que o *Tractatus* contenha teses acerca de uma realidade independente do sujeito que a representa.

É preciso, porém, mencionar alguns pontos em que o livro de McGinn deixa a desejar, embora, segundo penso, não acarretem uma dificuldade incontornável para seu argumento como um todo. Em primeiro lugar, apesar de, no primeiro capítulo, McGinn situar sua leitura entre duas interpretações extremas e opostas, à medida que o livro avança seus ataques passam a concentrar-se quase exclusivamente na leitura realista. Não se menciona quais são os argumentos específicos da leitura resoluto que ela considera inválidos, nem por que. Pode-se concluir, então, que ela parte do pressuposto de que era efetivamente a intenção de Wittgenstein apresentar descobertas positivas acerca do funcionamento da linguagem, e limita-se apenas a mostrar que essas descobertas não incluem teses metafísicas. A meu ver, é correto admitir, ao contrário da leitura resoluto, que os aforismos do *Tractatus* pretendem comunicar um conhecimento positivo acerca da lógica da nossa linguagem. Entretanto, ao apresentar sua interpretação como uma alternativa a duas outras leituras, esperar-se-ia que ela apresentasse argumentos específicos contra cada uma delas, o que efetivamente não é feito.

Outro ponto que seria preciso mencionar diz respeito à determinação do valor de verdade das proposições. Vimos que, apesar de recusar que objetos constituam um domínio independente da linguagem, McGinn não nega que o *Tractatus* aceite a existência de uma realidade independente que confere valor de verdade às proposições. Essa postura, a meu ver, é correta: o *Tractatus* parece defender uma teoria da verdade por correspondência, na qual o valor de verdade de uma proposição depende da existência ou inexistência de algo que está *fora* da linguagem. Isto é, ainda que se possa admitir que o sentido da proposição possa determinar-se *internamente* à linguagem, seu valor de verdade deve depender de um domínio externo e independente. Em uma leitura realista, isso pode ser entendido sem maiores dificuldades. Na medida em que objetos são concebidos como entidades externas e reais, eles estariam em posição de constituir esse domínio: proposições seriam verdadeiras ou falsas devido à existência ou

inexistência de algo absolutamente preciso, a saber, uma configuração possível desses objetos. Já na leitura de McGinn é mais difícil compreender de que maneira o valor de verdade se determinaria. Ele deve certamente depender de algo extralinguístico, mas nessa interpretação, como vimos, objetos não são entidades independentes da linguagem, já que se definem apenas pelas propriedades lógico-sintáticas do simbolismo. McGinn parece não enxergar uma grande dificuldade nisso, pois não se ocupa em demonstrar a possibilidade de uma articulação entre a linguagem, concebida como um domínio autônomo no que diz respeito à sua forma, e uma realidade independente que confere valor de verdade às proposições. Seria necessário mostrar que a recusa de um isomorfismo entre os dois domínios não torna absolutamente impossível uma comunicação entre eles. Isso, como foi dito, não é feito por McGinn em seu livro.

Não se deve pensar, porém, que essas dificuldades sejam de algum modo incontornáveis ou fatais para sua interpretação. Sem entrar na questão de se sua leitura é ou não correta ou definitiva, o livro de McGinn aborda de uma maneira detalhada todos os grandes temas do *Tractatus*, e captura alguns pontos que muitas vezes escapam a outros intérpretes. Trata-se, portanto, de uma grande contribuição à literatura secundária de Wittgenstein.

O Princípio de não-contradição e sua tradução para a aritmética transreal (The Principle of non-contradiction and its translation into transreal arithmetic)

Walter Gomide

Universidade Federal de Mato Grosso

Introdução

Não há dúvida de que a lógica contemporânea colocou sob análise ou suspeita um dos pilares do pensamento ocidental: o princípio de não-contradição. Desde o surgimento da lógica paraconsistente, a partir dos trabalhos do lógico brasileiro Newton da Costa, nos anos 60 do século XX, até o desenvolvimento do “dialeteísmo”, vertente filosófica capitaneada pelo lógico inglês Graham Priest, que postula que há de fato, no mundo, contradições verdadeiras, o princípio de não-contradição deixou de ser uma diretriz do pensamento racional, passando a ser, dentro de certos limites, atenuado (sobre a lógica paraconsistente, ver Da Costa [1993] e sobre o Dialeteísmo, ver Priest [2008])

Mas o que, de fato, é o princípio de não-contradição? Este artigo visa dar uma resposta a esta pergunta, de tal forma que instrumentos da aritmética transreal sejam usados a fim de oferecer uma tradução ilustrativa do comportamento do princípio de não-contradição, seja em sistemas clássicos, que adotam tal princípio como verdadeiro, seja nos ditos sistemas não-clássicos, que atenuam ou eliminam a verdade de tal princípio.

O princípio de não-contradição será tratado dentro da estrutura dos números transreais, novo domínio aritmético criado pelo Prof. James Anderson (ver Anderson, [2008]).

Os números transreais

Os números transreais consistem em uma generalização dos números reais **R**. Basicamente, a novidade introduzida pelos transreais é a permissão da divisão por zero, o que não é aceito no âmbito dos números reais. De fato, a introdução da divisão

irrestrita por zero, como definida nos transreais, não foi aceita de forma consensual na comunidade matemática. Em especial, a divisão de zero por zero (a qual dá origem ao número transreal “nullity”) é vista como uma mera tradução (ou “renomeação”) do já conhecido *NaN* (“Not a Number”), o qual aponta simplesmente que foi cometido um erro sintático na manipulação das operações aritméticas.

Para obter os números transreais \mathbf{R}^T dos números reais, primeiramente são introduzidos os seguintes números (ver Anderson, [2008]):

$$(1) \quad 1/0 = \infty;$$

$$(2) \quad -1/0 = -\infty.$$

Desta forma, chegamos ao domínio dos *números reais estendidos*:

$$(3) \quad \mathbf{R}^E = \mathbf{R} \cup \{-\infty, \infty\}.$$

A partir dos números reais estendidos, chega-se ao números transreais mediante a introdução da constante Φ , definida da forma seguinte:

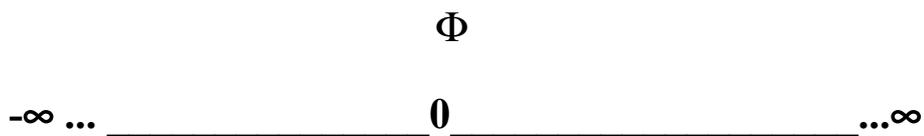
$$(4) \quad 0/0 = \Phi.$$

Desta forma, os transreais \mathbf{R}^T são definidos como:

$$(5) \quad \mathbf{R}^T = \mathbf{R}^E \cup \{\Phi\}.$$

É possível definir as operações usuais da aritmética nos números transreais de forma consistente (ver ANDERSON *et al*, [2007]).

Os números transreais podem ser visualizados da forma seguinte:



A constante Φ (denominda por Anderson de *nullity*) ocupa uma posição *fora* da reta representativa dos números reais estendidos $\mathbf{R}^E = \mathbf{R} \cup \{-\infty, \infty\}$.

Um resultado interessante se depreende no que diz respeito à relação de ordem $<$ que se estabelece nos transreais \mathbf{R}^T . Para qualquer número real r , valem sempre as seguintes desigualdades:

$$(6) \quad r < \infty;$$

$$(7) \quad -\infty < r$$

Entretanto, a constante Φ não é ordenada em \mathbb{R}^T , isto é, para qualquer número $t \in \mathbb{R}^T$, temos sempre que:

$$(8) \quad \neg(t < \Phi);$$

$$(9) \quad \neg(\Phi > t).$$

O Princípio de não-contradição “traduzido” para a aritmética transreal

Consideremos, inicialmente, variáveis proposicionais ψ e ϕ que tomam por valores constantes de predicado de aridade 0 – letras sentencias, que pertencem ao vocabulário de uma linguagem L . Suponhamos agora que há uma função ν , definida no conjunto $P = \{\psi, \phi\}$, e que toma por valores números transreais não-negativos. Consideremos também que há uma função σ que associa os valores transreais de ν com elementos do seguinte conjunto Λ :

$$\Lambda = \{T, F, c, i\},$$

Λ é o conjunto de todos os possíveis valores de verdade que uma constante de predicado, uma vez interpretada, pode ter, isto é: “T” é o *verdadeiro*, “F” é o *falso*, “c” é *tanto verdadeiro como falso*, e “i” é *nem verdadeiro, nem falso*.

A partir da função ν (cujos valores nos números transreais maiores ou igual a zero, a partir do conjunto Λ , são definidos arbitrariamente), podemos definir uma função σ que associa os valores transreais não-negativos (arbitrários) de ν no domínio Λ :

- a) Se $\nu(Y_{\in P}) = r_{\in \mathbb{R}(\text{reais}) > 0}$, então $\sigma(\nu(Y_{\in P})) = c$;
- b) Se $\nu(Y_{\in P}) = \infty$, então $\sigma(\nu(Y_{\in P})) = T$;
- c) Se $\nu(Y_{\in P}) = 0$, então $\sigma(\nu(Y_{\in P})) = F$;
- d) Se $\nu(Y_{\in P}) = \Phi$, então $\sigma(\nu(Y_{\in P})) = i$.

Claramente, a função σ atribui valores de verdade a partir de valores transreais que são imagens de parâmetros proposicionais de uma linguagem L .

Consideremos agora que L , além de constantes de predicado, também possui os símbolos lógicos usuais do alfabeto do cálculo proposicional, isto é, $\lambda = \{\neg, \rightarrow, \leftrightarrow, \wedge, \vee\}$ é um subconjunto de L . Sendo assim, em L , também podem se definir fórmulas

moleculares, resultantes de conhecidas regras recursivas de formação de fórmulas (ver PRIEST, p.4, [2008]).

A linguagem L pode ser interpretada. De fato, a função σ tem o papel de interpretar (isto é, atribuir um valor de verdade) às constantes de predicado presentes em L (em um primeiro momento, considera-se aqui que tais constantes de predicado são letras sentenciais, isto é, constantes de predicado que são interpretadas não como conjuntos, mas diretamente como valores de verdade). Uma vez sendo interpretadas as constantes de predicado, as fórmulas também admitem ser interpretadas por meio de conhecidas regras semânticas que “espelham” os procedimentos recursivos utilizados para a formação de fórmulas moleculares. Na semântica das lógicas ditas clássicas, de natureza bivalente, as regras pelas quais as fórmulas complexas são valoradas a partir de seus componentes atômicas são as seguintes (no que se segue, ψ e ϕ são parâmetros proposicionais que percorrem o domínio das constantes de predicado de L, e σ a função anteriormente definida):

- 1) $\sigma(v(\neg\psi)) = T$ sse $\sigma(v(\psi)) = F$.
- 2) $\sigma(v(\psi \wedge \phi)) = T$ sse $\sigma(v(\psi)) = T$ e $\sigma(v(\phi)) = T$.
- 3) $\sigma(v(\psi \vee \phi)) = T$ sse $\sigma(v(\psi)) = T$ ou $\sigma(v(\phi)) = T$.
- 4) $\sigma(v(\psi \rightarrow \phi)) = T$ sse $\sigma(v(\psi)) = F$ ou $\sigma(v(\phi)) = T$
- 5) $\sigma(v(\psi \leftrightarrow \phi)) = T$ sse $\sigma(v(\psi)) = \sigma(v(\phi))$

É sabido que há um único operador lógico, a barra de Sheffer “ \uparrow ”, por meio do qual todas as condições valorativas acima expressas podem ser equivalentemente definidas só com tal conectivo. Mais especificamente, os valores de verdade que podem ser atribuídos à barra de Sheffer obedecem à seguinte equação:

$$6) \sigma(v(\psi \uparrow \phi)) = \sigma(v(\neg(\psi \wedge \phi))).$$

Segue-se, então, que:

$$7) \sigma(v(\psi \uparrow \psi)) = \sigma(v(\neg(\psi \wedge \psi))) = \sigma(v(\neg\psi)).$$

Desta forma,

$$8) \sigma(v(\psi \uparrow (\psi \uparrow \psi))) = \sigma(v(\neg(\psi \wedge \neg\psi)))$$

A expressão acima nos dá a maneira pela qual o princípio de não-contradição, sob o ponto de vista semântico, pode ser apresentado somente com a barra de Sheffer.

Cabe notar aqui que a barra de Sheffer admite uma tradução para a aritmética transreal. De fato, a expressão:

$$9) \nu(\psi \uparrow \varphi) = (1/\nu(\psi) + 1/\nu(\varphi))$$

“mapeia” a barra de Sheffer para o domínio da aritmética transreal. Uma vez que na aritmética transreal $1/0 = \infty$ e $1/\infty = 0$, então fica evidente que os valores admitidos pela expressão acima, para quaisquer valores transreais não-negativos de $\nu(\psi)$ e $\nu(\varphi)$, serão os valores semanticamente esperados para a barra de Sheffer, isto é¹:

$$10) \sigma(\nu(\psi \uparrow \varphi)) = \sigma((1/\nu(\psi) + 1/\nu(\varphi))).$$

O Princípio de não-contradição, expresso por meio da barra de Sheffer, pode ser traduzido à aritmética transreal através de:

$$11) \nu(\psi \uparrow (\psi \uparrow \psi)) = (1/\nu(\psi) + 1/\nu(\psi \uparrow \psi)),$$

o que nos dá:

$$12) \sigma(\nu(\psi \uparrow (\psi \uparrow \psi))) = \sigma((1/\nu(\psi) + 1/\nu(\psi \uparrow \psi))).$$

Chamemos $(1/\nu(\psi) + 1/\nu(\psi \uparrow \psi))$ de π_ψ . Temos então o seguinte:

Se $\pi_\psi \neq \infty$, então $\sigma(\pi_\psi) \neq T$ (o princípio de não-contradição, semanticamente, se comporta de forma *não-clássica*). Por sua vez, se $\pi_\psi = \infty$, então $\sigma(\pi_\psi) = T$ (o princípio de não-contradição, semanticamente, se comporta de forma *clássica*).

Portanto, a grandeza π_ψ , inteiramente definida na aritmética transreal, serve como um parâmetro do comportamento semântico do princípio de não-contradição. Mas isto nos leva a um problema mais fundamental:

A equação acima, 10), nos dá uma tradução para a aritmética transreal da barra de Sheffer. A partir desta tradução, uma versão (trans) aritmética do princípio de não-contradição surge através da grandeza π_ψ . Mas o que, de fato, significa tal grandeza, a ponto de ser um parâmetro perfeitamente adequado para nos indicar o comportamento semântico do princípio de não-contradição?

Para responder à pergunta acima formulada, duas hipóteses são introduzidas. A primeira, busca resolver a questão mediante uma interpretação (trans) métrica² da barra de Sheffer; a segunda, associa a barra de Sheffer à medidas de conjuntos

¹ Na aritmética transreal, se $1/\nu(\psi)$ é igual a Φ ou se $1/\nu(\varphi)$ é igual a Φ , então $(1/\nu(\psi) + 1/\nu(\varphi))$ é igual a Φ , para qualquer valor que a parcela restante possa ter; se tanto $1/\nu(\psi)$ for um número real e $1/\nu(\varphi)$ também for um número real, então $(1/\nu(\psi) + 1/\nu(\varphi))$ é um número real, como era de se esperar. Se uma das parcelas envolvidas for um número real e a outra for um transreal pertencente a $\{\infty, \Phi\}$, então a soma nos dá um transreal, Φ ou ∞ .

² Por “interpretação (trans) métrica” entende-se a compreensão da barra de Sheffer como distância, no espaço (trans) métrico, entre conjuntos. Vale dizer que tal distância toma por valores tantos números reais como Φ ou ∞ .

A barra de Sheffer como distância entre conjuntos

Consideremos os predicados P e Q pertencentes a uma linguagem L . A estes dois predicados, mediante uma função f , façamos corresponder duas regiões Π e K de um espaço trans-métrico G . Por um espaço trans-métrico, entende-se uma generalização da noção de espaço métrico, de tal forma que em G estão definidas tanto distâncias de valores reais, quanto valores trans-reais. Os postulados que definem um espaço trans-métrico são os seguintes (ANDERSON, [2008]):

- $t(a,b) = t(b,a)$.
- $\neg(t(a,b) < 0)$
- $t(a,b) = 0$ sse $a = b$
- $\neg(t(a,b) + t(b,c) < t(a,c))$

Nos postulados acima, a função t é chamada de trans-métrica, e pode tomar por valores tanto números reais quanto transreais.

Seja agora a função f que associa os predicados P e Q a regiões do espaço transmétrico G .

- $P \rightarrow^f \Pi$.
- $Q \rightarrow^f K$.

Consideremos agora que se possa avaliar a distância entre Π e K (uma maneira de se fazer isto seria por meio da noção de “distância de Hausdorff” entre conjuntos³).

Desta maneira, a barra de Sheffer, traduzida na aritmética transreal, poderia ser vista como uma função g desta distância entre regiões no espaço trans-métrico, isto é:

- $v(P \uparrow Q) = g[t(\Pi, K)]$,

de tal forma que

- se $t(\Pi, K) = \infty$, então $v(P \uparrow Q) = \infty$;
- se $t(\Pi, K) \neq \infty$, então $v(P \uparrow Q) \neq \infty$.

³ A distância de Hausdorff entre dois subconjuntos X e Y de um determinado espaço métrico definida da forma seguinte: $D_H(X, Y) = \max \{ \sup_{x \in X} \inf_{y \in Y} d(x, y), \sup_{y \in Y} \inf_{x \in X} d(x, y) \}$.

A distância infinita entre conjuntos pode ser obtida da forma seguinte. Fazemos os predicados P e Q serem mapeados, no espaço transmétrico G , pelos conjuntos $\Pi = \{\infty\}$ e $K = \{c\}$. Definamos a transmétrica de G como:

$$t(a,b) = |a - b|.$$

No caso em questão, $t(\infty,c) = \infty^4$. Desta maneira, se tomarmos a distância de Hausdorff entre os conjuntos Π e K , adaptada ao espaço transmétrico (ver nota 3), então $t(\Pi,K) = \infty$.

Portanto, sendo Π^c a região complementar a Π no espaço transmétrico G , podemos introduzir um significado métrico para a tradução π_ψ do princípio de não contradição para a aritmética transreal:

- se $t(\Pi, \Pi^c) = \infty$, então $\pi_\psi = \infty$ (isto é, $\sigma[(\pi_\psi)] = T$);
- se $t(\Pi, \Pi^c) \neq \infty$, então $\pi_\psi \neq \infty$ (isto é, $\sigma[(\pi_\psi)] \neq T$).

As condições acima j) e k) nos dão um critério métrico para que o princípio de não contradição seja verdadeiro: a distância entre as traduções transmétricas de P e não- P (respectivamente, Π e Π^c) deve ser infinita; caso contrário, o princípio de não-contradição tem valor diferente do verdadeiro.

Assim, j) e k) nos oferecem um critério inteiramente métrico para avaliar o comportamento semântico do princípio de não-contradição.

A barra de Sheffer como medida de conjuntos no espaço transmétrico

Consideremos que os predicados P e Q sejam traduzidos para o espaço transmétrico G , por meio de uma função f , como as regiões Π e K :

$$a) P \rightarrow \Pi$$

$$b) Q \rightarrow K.$$

Como se viu anteriormente,

⁴ Na aritmética transreal, valem as seguintes identidades: para todo número real r , $|\infty - r| = |r - \infty| = \infty$.

$$c) \nu(P \uparrow Q) = \nu(\neg(P \wedge Q)) = \nu(\neg P \vee \neg Q).$$

Consideremos agora que todos os operadores lógicos, por meio de uma função γ , sejam traduzidos para os seus correlatos conjuntísticos (a existência de tal função é garantida pelo fato de que, tanto o cálculo proposicional e a teoria dos conjuntos são uma álgebra booleana e, como tal, são intertraduzíveis). Desta forma, pode-se estipular uma função μ que associa conjuntos aos universos dos transreais:

$$d) \quad \mu(\gamma(P \uparrow Q)) = \mu(\gamma(\neg(P \wedge Q))) = \mu(\gamma(\neg P \vee \neg Q)) = \mu((\Pi \cap K)^c) = \mu(\Pi^c \cup K^c)$$

Seja L uma linguagem lógica na qual se definem expressões do cálculo proposicional e seja L_c uma linguagem na qual se definem as operações usuais da teoria dos conjuntos. Desta maneira, pode-se estabelecer a seguinte composição das funções ν , γ e μ , sendo dadas as condições seguintes e), f), e g):

$$e) L \rightarrow^{\nu} R^T$$

$$f) L \rightarrow^{\gamma} L_c$$

$$g) L_c \rightarrow^{\mu} R^T$$

Portanto,

$$h) L \rightarrow^{\nu=\mu\gamma} R^T.$$

Logo,

$$i) \nu(P \uparrow Q) = \mu\gamma(P \uparrow Q) = \mu(\Pi^c \cup K^c) = 1/\mu(\Pi^c) + 1/\mu(K^c).$$

Portanto,

$$j) \mu(P \uparrow (P \uparrow P)) = 1/\mu(\Pi^c) + 1/\mu((\Pi^c)^c) = 1/\mu(\Pi^c) + 1/\mu(\Pi).$$

Desta forma, sendo Π^c o conjunto complementar de Π em relação ao espaço métrico G , podemos considerar que $\pi_p = \mu(\Pi \cup \Pi^c)$ é a medida de dois conjuntos complementares em relação a um espaço trans-métrico G – facilmente se constata que a função μ é uma função de medida, posto que satisfaz os postulados que definem uma função como “medida”, a saber:

Seja f uma função definida em uma σ -álgebra⁵, com valores entre $[0, \infty]$ (os valores de função poderiam estar no conjunto $[0, \infty] \cup \{\Phi\}$), f é uma função-medida se satisfizer os seguintes postulados:

$$M1) f(\emptyset) = 0;$$

$$M2) f(E_1 \cup E_2 \cup E_3 \dots) = f(E_1) + f(E_2) + f(E_3) + \dots$$

Sendo assim, a grandeza π_p mede o “tamanho coberto” pela região Π e sua complementar Π^c . Assim, institui-se um critério de medida para que o princípio de não-contradição seja verdadeiro: o tamanho coberto pelas traduções transmétricas de P e não- P deve ser infinito, isto é, o conjunto formado pela união das traduções transmétricas de P e não- P deve ter medida infinita, caso contrário, o princípio de não-contradição não é verdadeiro.

Referências Bibliográficas

- 1) Anderson, J. A.D. W. “Representing Geometrical Knowledge”. *W. Phil. Trans.R.Soc.* London, series B, Vol. 352, no 1358, pp 1129-1139;
- 2) _____. “Perspex Machine XI: Topology of Transreal Numbers”. In: JMEC5 2008, S.1., Hong Kong, pp 330-338;
- 3) Bell, J. L., De Vidi, d. and Soloman, G (2001), *Logical Options: An Introduction to Classical and Alternative Logics*. (Petersborough, ON: Broadview Press);
- 4) Chaitin, G. *The Limits of Mathematics*, Spring Verlag, 1988;
- 5) _____. *Meta Math! The Quest for Omega*, Pantheon Books, 2005;
- 6) Chang, C.C. (1963), ‘Logic with Positive and Negative Truth Values’, *Acta Philosophica Fennica* 14: 19-39
- 7) Da Costa, N. C. A. (1974), ‘On the Theory of Inconsistent Formal Systems’, *Notre Dame Journal of Formal Logic* 15: 497-510;
- 8) _____. (1993), *Sistemas Formais Inconsistentes*, Curitiba: Editora UFPR
- 9) Dória F.A. and J.F. Costa. “Special Issue on Hypercomputation”, *Applied Mathematics and Computation*, Vol. 178, 2006;
- 10) Dummett, M. (1959), *Elements of Intuitionism*, (Oxford: Oxford University Press);
- 11) Fine, K. (1975), ‘Vagueness, Truth and Logic’, *Synthese* 30: 265-300; reeditado em Keefe and Smith (1996), ch 9;
- 12) Gabbay, D. and Guenther, F. (1983-1989), (eds), *Handbook of Philosophical Logic* (Dordrecht: Kluwer Academic Publishers);
- 13) _____. (2001-), *Handbook of Philosophical Logic*, second edition (Dordrecht: Kluwer Academic Publishers);

⁵ Uma σ -álgebra em X satisfaz os seguintes postulados: (1) \emptyset está em X ; (2) Se $E_1, E_2, \dots, E_n, \dots$ estão em X , então $E_1 \cup E_2 \dots \cup E_n \cup \dots$ está em X .

- 14) Gödel, K. (1930), 'Die Vollständigkeit der Axiome des logischen Funktionenkalküls', *Monatshefte für Mathematik und Physik* 28: 173-98;
- 15) Gottwald, S. (2001), *A Treatise on Many Valued Logics* (Baldock: Research Studies Press Ltd);
- 16) Haack, S. (1974), *Deviant Logic* (Cambridge: Cambridge University Press);
- 17) Hájek, P. (1998), *Metamathematics of Fuzzy Logic* (Dordrecht: Kluwer);
- 18) Heijenoort, J. van (1967), *From Frege to Gödel* (Cambridge, MA: Harvard University Press);
- 19) Jaskowski, S. (1969), 'Propositional Calculus for Contradictory Deductive Systems', *Studia Logica* 24: 143-57;
- 20) Kleene, S.C. (1952), *Introduction to Metamathematics* (Amsterdam: North Holland);
- 21) Malinowski, G. (1993), *Many-Valued Logics* (Oxford: Oxford University Press);
- 22) Priest, G. *An Introduction to Non-Classical Logic. From If to Is*. Second Edition, Cambridge University Press, 2008;

Como os nomes próprios realmente referem

Claudio Ferreira Costa ¹

Universidade Federal do Rio Grande do Norte

Resumo: Nesse artigo é sugerido o esboço de uma versão aprimorada da tradicional teoria descritivista dos nomes próprios. Essa versão é capaz de explicar o conteúdo informativo dos nomes próprios e o seu contraste, como designadores rígidos, com as descrições definidas, além de responder de forma mais adequada aos contraexemplos usualmente apresentados contra o descritivismo.

Palavras-chave: descritivismo, nomes próprios, referência.

Abstract: This paper contains a sketch of a more developed descriptivist theory of proper names. This version is able to explain the informative content of proper names and why they are rigid designators by contrast with definite descriptions. Moreover, it is able to answer in a more convincing way the counterexamples usually presented against descriptivism.

Key words: descriptivism, proper names, reference.

Qual é o mecanismo pelo qual nomes próprios como “Aristóteles”, “Paris”, “Vênus” etc. são capazes de designar seus portadores? A resposta mais antiga é descritivista. Segundo a teoria do agregado de descrições, que foi sugerida nos escritos de Frege, Russell e Wittgenstein, e que encontra a sua exposição última em um artigo de John Searle,² o que alguém tem em mente com um nome próprio de maneira a ser capaz de usá-lo referencialmente é exprimível por *um subconjunto indefinido de um conjunto aberto de descrições co-referenciais* (minimamente, um subconjunto contendo uma única descrição).³ Assim, um nome próprio como “Aristóteles” pode vir no lugar de

¹ Doutor em filosofia pela universidade de Konstanz, professor do Departamento de Filosofia da UFRN, bolsista de pesquisa individual do CNPq, e-mail: ruvstof@gmail.com.

² J. R. Searle: “Proper Names”.

³ Essa é a formulação sinóptica da versão searleana apresentada por Susan Haack em *Philosophy of Logics*, p. 58.

descrições definidas (geralmente iniciadas com um artigo definido) como “o estagirita”, “o autor da *Ética a Nicômaco*”, “o autor da *Metafísica*”, “o discípulo de Platão”, “o fundador do Liceu”, “o tutor de Alexandre”.

Para Saul Kripke, um problema com as teorias descritivistas é que, embora um nome próprio se aplique necessariamente ao seu portador em qualquer circunstância na qual ele exista, esse nome pode ser aplicado ao seu portador sem que nenhuma das descrições definidas a ele usualmente associadas necessariamente se aplique.⁴ Assim, podemos imaginar um mundo possível no qual Aristóteles existiu, mas morreu ainda criança, não tendo sido discípulo de Platão e nem escrito nenhuma das obras a ele atribuídas. E também podemos imaginar um mundo possível no qual Aristóteles existiu, mas não nasceu em Estagira e sim em Roma, trezentos anos mais tarde. Fica assim claro que nenhuma das descrições que associamos ao nome próprio se aplica necessariamente. Além do mais, Kripke notou que uma pessoa pode usar um nome próprio referencialmente, mesmo tendo em mente uma única descrição, que pode ser indefinida ou até mesmo incorreta. Assim, uma pessoa pode se referir a Feynman, dele sabendo apenas que foi *um* cientista norte-americano, e alguém pode perfeitamente se referir a Einstein pensando incorretamente que ele foi o inventor da bomba atômica. Considere ainda o caso de nomes de personagens semifictícios, como “Robin Hood”. Sabemos que deve ter existido alguém que esteve na origem desse personagem, mas nada sabemos sobre ele, nem mesmo se foi realmente um fora da lei ou se ele chamava Robin Hood! Em nenhum desses casos a descrição adquire um papel relevante.

A solução encontrada por Kripke, Keith Donnellan, Michael Devitt e outros, foi externalista e causal. Esses filósofos concluíram que aquilo que suporta a referência de um nome próprio é uma cadeia causal-histórica externa, que para Kripke começa com a primeira denominação do objeto através do nome, o seu “batismo”. Se eu profiro o nome “Aristóteles” e esse for o último elo de uma imensamente longa e complexa cadeia causal-histórica que começou com o batismo de Aristóteles em 384 a.C. em Estagira, isso é suficiente para eu me referir a Aristóteles. Descrições podem

⁴ Ver Saul Kripke: *Naming and Necessity* cap. II. As outras objeções de Kripke se encontram todas no mesmo livro. Para uma resposta relevante, curiosamente passada em silêncio pelos defensores da concepção causal-histórica, ver J. R. Searle: *Intentionality: an Essay in the Philosophy of Mind*, cap. 9.

acompanhar o meu uso do nome “Aristóteles”, mas a sua função será meramente auxiliar. Embora essa hipótese possa a primeira vista parecer fantástica, há muito ela se tornou a nova ortodoxia.

Dois tipos fundamentais de descrições

A resposta internalista que eu desejo sugerir aqui consiste na defesa de uma forma mais sofisticada de descritivismo. Ela se resume na descoberta de uma meta-regra cuja função é a de selecionar elementos do agregado de descrições, provendo-lhe de uma estrutura valorativa adequada. É só quando o descritivismo é acrescido dessa meta-regra, exprimível na forma de uma meta-descrição, que ele ganha o poder explicativo que merece, tornando a hipótese causal-histórica dispensável como meio de explicar a referência.

Para chegarmos onde queremos precisamos primeiramente distinguir as descrições *fundamentais* das *descrições auxiliares*. As descrições fundamentais expressam critérios primários de identificação, enquanto as descrições auxiliares várias apenas nos remetem às descrições fundamentais na busca do portador do nome, merecendo nesse aspecto serem descartadas.

Esse ponto pode ser primeiramente mostrado através de exemplos de descrições tipicamente auxiliares. Considere o caso de descrições definidas como “o tutor de Alexandre”, “o fundador do Liceu”, “o maior discípulo de Platão”, todas elas concernentes a Aristóteles. Elas são descrições definidas usuais, mas perfeitamente dispensáveis: afinal, Aristóteles continuaria sendo ele mesmo, ainda que não tivesse sido nada disso. Outro caso é o de descrições definidas pouco usuais e dispensáveis como, no que concerne a Aristóteles, “o neto de Achaeon”, “o amante de Herphylis” e “a pessoa que viajou com Teofrasto à ilha de Lesbos”. Um terceiro caso é o de descrições definidas como “o estagirita” e “o mestre dos que sabem” com respeito a Aristóteles. Essas descrições tem função expressiva ou metafórica, cumprindo com um papel mnemônico. Mas parece certo que como tais elas pouco fazem para caucionar a identificação do objeto referido. Finalmente, existem descrições que são adventícias e temporárias, como a usada pelo aluno que só sabe dizer de Aristóteles que ele foi “o

filósofo mencionado pelo professor”. Chamo a todas essas descrições de auxiliares porque, como veremos, apesar de nos orientarem na identificação do portador do nome e apesar de constituírem parte secundária de seu conteúdo informativo, elas terminam sendo prescindíveis.

Minha sugestão é a de que as descrições fundamentais para a referência do nome próprio são de outro tipo e que as formas tradicionais da teoria do agregado são enganosas, em parte por seus proponentes terem sido desviados do que é relevante ao escolherem como exemplos descrições auxiliares, como “o tutor de Alexandre e o fundador do Liceu” para Aristóteles (Frege) e “a criança retirada do Nilo pela filha do faraó” para Moisés (Wittgenstein).

Minha proposta é a de que as descrições realmente fundamentais são expressões lingüísticas de duas espécies de regras identificadoras do objeto, que são:

- A. REGRA LOCALIZADORA: que estabelece a *localização e carreira espacio-temporal* do portador do nome próprio,
- B. REGRA CARACTERIZADORA: que estabelece uma *caracterização* daquilo que consideramos como *mais relevante* no portador do nome próprio, de modo a justificar nossa aplicação do mesmo.

Assim, para um nome próprio como “Aristóteles” a descrição “a pessoa nascida em Estagira em 384 a.C., que viveu grande parte da sua vida em Atenas e que faleceu em Chalcis em 322 a.C.” exprime resumidamente nossa regra localizadora de Aristóteles no espaço e no tempo. Já a descrição que permite caracterizar Aristóteles por aquilo que nele consideramos importante de modo a justificar que ele seja identificado através desse nome próprio pode ser resumida como “o autor do conteúdo relevante do *opus* aristotélico”.

Mas o que justifica minha aposta nas regras localizadora e caracterizadora como sendo as regras-descrições fundamentais? A resposta consiste simplesmente no apelo às intuições oferecidas por nossa linguagem natural. Em atenção a isso J. L. Austin, o filósofo da linguagem ordinária, aconselhava-nos o uso dos melhores dicionários como um método para distinguirmos distinções semânticas que pudessem importar ou não importar filosoficamente. Como nomes próprios raramente são dicionarizados falta-nos esse recurso. Mas como eles são muitas vezes enciclopedizados, essa falta é

sobejamente compensada quando procuramos as condições de referência de nomes próprios oferecidas em enciclopédias. Com efeito, as enciclopédias geralmente explicam o que os nomes próprios querem dizer partindo das descrições fundamentais e não de descrições auxiliares. Eis o que encontro sobre o nome “Aristóteles” em meu *Penguin Dictionary of Philosophy*, que escolho por ser o mais conciso:

(384 - 322 a.C.) Nascido em Estagira no norte da Grécia. Aristóteles produziu o mais completo e poderoso sistema filosófico da antiguidade. (Segue-se uma breve exposição da vida de Aristóteles, seguida de um resumo das principais obras...)

Essa descrição concentra-se nos critérios identificadores dos tipos A e B. Se consultarmos a elucidação lexical de outros nomes próprios, não só de pessoas, mas também de coisas, como “Taj Mahal”, “Paris”, “China” e “Rio Amazonas”, encontraremos resultados semelhantes.

Para tornar mais claro o caráter fundamental das descrições localizadora e caracterizadora, tente imaginar um nome próprio usual sem que suas descrições fundamentais tenham aplicação alguma. Imagine, por exemplo, um Aristóteles que nunca teve nada a ver com a filosofia ou com a ciência, que não nasceu na Grécia antiga e que teve uma localização, carreira e origem espacio-temporal completamente diversa daquela tida pelo Aristóteles filósofo. Imagine, glosando um exemplo de John Searle, que a pessoa chamada pelo nome “Aristóteles” tenha sido apenas um vendedor de peixes veneziano que viveu na Renascença tardia e que nunca teve nada a ver com a filosofia... Certamente não o reconheceremos como sendo o nosso Aristóteles, mas alguma outra pessoa com o mesmo nome!⁵

Essa observação se complementa com outra: a de que na completa ausência de aplicação das descrições fundamentais as descrições auxiliares deixam de ser úteis. Para que esse ponto se evidencie tente imaginar que as descrições auxiliares que geralmente associamos a Aristóteles se apliquem ao nosso vendedor de peixes veneziano: descobrimos que ele ensinou alguém chamado “Alexandre”, que ele fundou um Liceu e foi chamado de “o mestre dos que sabem”. Nenhuma dessas estranhas coincidências

⁵ J. R. Searle: “Proper Names and Descriptions”, p. 490.

seria capaz de nos fazer admitir que o vendedor de peixes seja realmente aquela pessoa a que pretendemos nos referir com o nome “Aristóteles”. Afinal, o Alexandre que ele ensinou não pode ter sido o maior conquistador de todos os tempos, não é possível que o vendedor de peixes iletrado tenha verdadeiramente fundado o Liceu aristotélico, nem que o apelido “o mestre dos que sabem” seja o mesmo que foi usado por Dante na *Divina Comédia*. Pode ser que se descubra então que esse Aristóteles vendedor de peixes foi um fanfarrão que se acreditava sábio, que o Alexandre que ele ensinou foi um pescador, que o Liceu foi o nome que as pessoas deram ao local onde ele reunia as pessoas para contar as suas estórias, e que a descrição “o mestre dos que sabem” tenha sido ironicamente usada por seus conhecidos... Mesmo que todas as descrições auxiliares fossem aplicáveis, na falta da aplicação das descrições fundamentais elas não seriam capazes de produzir mais do que uma sensação de estranha coincidência, de persiflage do real. A conclusão parece ser a seguinte: descrições auxiliares só serão capazes de contribuir para a referência se forem articuladas dentro de um contexto definido pelas próprias descrições fundamentais para as quais elas nos direcionam, tornando-se inúteis na ausência disso.

A regra de identificação do nome próprio

Uma vez que encontramos as descrições fundamentais, a questão seguinte é saber como elas são exigidas para a aplicação de um nome próprio qualquer.

Uma primeira consideração a ser feita é que para a identificação do portador do nome próprio a satisfação de “A & B”, ou seja, da conjunção das condições, é desnecessária. Há razões conclusivas para se pensar assim. A primeira é que há nomes próprios que constitutivamente possuem apenas uma regra caracterizadora. Considere o nome “Universo”. A descrição caracterizadora é “tudo o que existe”. Mas pelo próprio fato de ser tudo o que existe (existiu e existirá), o universo não se encontra nem no espaço nem no tempo. Há também nomes próprios que só possuem regra localizadora. Digamos que o centro de um dado círculo seja casualmente denominado “Z”. Aqui a localização é o que importa, pois não há outro motivo para a sua consideração.

Outra razão para se pensar que a conjunção “A & B” é desnecessária vem da consideração de situações contrafactuais. Imagine, pois, um mundo possível *m1*, muito próximo ao nosso, no qual Aristóteles nasceu em Estagira em 384 a.C., filho de Nicômaco, o médico da corte de Felipe, mas que ele morreu de febre aos dezessete anos, em sua viagem para Atenas, não chegando a escrever o *opus* aristotélico. Nesse caso admitiremos talvez que nosso Aristóteles “em potência” existiu em *m1*. Nesse caso apenas a regra localizadora é aplicada, mas o nome próprio “Aristóteles” encontra a sua designação. Além disso, podemos conceber um mundo possível *m2*, também muito próximo ao nosso, no qual Aristóteles viveu em Roma cerca de trezentos anos mais tarde, tendo lá escrito o seu *opus*. Nesse caso tenderemos a dizer que *m2* também teve o seu Aristóteles, embora ele tenha existido em lugar e época diversos. Podemos até mesmo imaginar que os indivíduos aqui imaginados não se chamavam Aristóteles, pois descrições do tipo “a pessoa de nome N” também são auxiliares (se em um mundo possível o autor do *opus* aristotélico se chamasse Pitacus, reconheceríamos Pitacus como sendo o nosso Aristóteles).

Do fato de que as condições A e B não são isoladamente necessárias podemos concluir que nem a descrição localizadora nem a descrição caracterizadora são essenciais, se por essencial se entende algo que é necessário. Podemos, no entanto, conceber que a satisfação de uma disjunção “A ou B” das regras-descrições fundamentais seja uma condição minimamente capaz de dotar os nomes próprios de referência. O que efetivamente não parece possível, contudo, é que possamos imaginar que um nome próprio se aplique em um mundo possível em que “ $\sim A \ \& \ \sim B$ ” seja o caso, a dizer, em que nenhuma das regras-descrições fundamentais se aplique. Esse é o caso do exemplo já considerado de Searle, em que ele imagina um especialista em Aristóteles que veio nos dizer que descobriu que Aristóteles não foi nem grego nem filósofo, mas um obscuro vendedor de peixes veneziano que viveu na renascença tardia... O mesmo aconteceria se alguém nos dissesse que Aristóteles foi, na verdade, um armador grego que viveu no século XX, foi amante de Maria Callas e se casou com Jackeline. Afinal, Aristóteles Onassis não satisfaz nem a descrição localizadora nem a descrição caracterizadora para o estagirita.

Nesse ponto pode ser objetado que as condições A e B não precisam em si mesmas ser inteiramente satisfeitas. Um nome próprio pode se aplicar ao seu portador mesmo quando as descrições fundamentais se aplicam apenas parcialmente ou quando uma só se aplica e mesmo assim se aplica apenas parcialmente! Afinal, no mundo *m1*, no qual Aristóteles nasceu em Estagira em 384 a.C., mas morreu aos dezessete anos, não só a condição caracterizadora não está sendo satisfeita, mas a condição localizadora está sendo apenas parcialmente satisfeita, já que ele não teve a carreira espácio-temporal esperada: ele não viveu em Atenas, não visitou Lesbos e nem morreu em Chalcis em 322 a.C. Igualmente, no mundo possível *m2*, em que Aristóteles viveu em Roma mais de duzentos anos depois, podemos conceber que ele tenha escrito apenas a *Ética a Nicômano*, a *Metafísica* e alguns outros trabalhos menores. Se não houver nenhum Aristóteles grego para competir com ele, nós tenderemos a admitir Aristóteles existiu realmente em *m2*, mesmo que grande parte da condição B não esteja sendo satisfeita e coisa alguma da condição A tenha sido satisfeita. É fácil, porém, lidar com essa objeção. Basta exigir satisfação *suficiente* e não mais completa do termo ou dos termos da disjunção. Se o Aristóteles de *m1* morreu logo após o nascimento e se o Aristóteles de *m2* apenas escreveu alguma obra que é apenas improvavelmente atribuída a Aristóteles, começaremos a duvidar de que a pessoa em questão tenha sido nosso Aristóteles.

Finalmente, é necessário considerar o caso do mundo possível no qual existem dois ou mais objetos que satisfazem as condições fundamentais. Nesse caso o verdadeiro objeto de referência do nome próprio será aquele que tiver satisfeito as descrições fundamentais *de modo mais completo*. Se no mundo *m3* além do Aristóteles de Estagira tivesse existido um filósofo romano com o nome de Aristóteles que tivesse escrito o *opus* Aristotélico pouco antes da conquista da Grécia pelos romanos, nós veríamos nisso uma inexplicável coincidência. Mas preferiríamos considerar o Aristóteles de Estagira como sendo o nosso Aristóteles, uma vez que ele satisfaz a condição de localização, além da condição de caracterização.

Juntando as condições fundamentais e as condições adicionais recém-consideradas estamos em condições de estabelecer o que parece ser a forma de qualquer regra de identificação de nome próprio. Ela é estabelecida pelo que chamo de uma regra meta-descritiva, uma regra de regras, posto que as descrições pertencentes ao agregado

também são expressões de regras. Essa regra meta-descritiva, expressa por uma meta-descrição, tem a função de organizar os agregados de regras-descrições de qualquer nome próprio eventualmente dado. Ela pode ser apresentada como uma regra constituída por três condições, que chamaremos de (i) a *condição de significação*, (ii) a *condição de suficiência* e (iii) a *condição de predominância*. Ei-la:

RMD:

Um nome próprio N se aplica a um objeto da classe C

see

- (i) se condição A para *N* e/ou a condição B para *N* for(em) satisfeitas pelo objeto,
- (ii) se a satisfação dessa ou dessas condição/condições por esse objeto for suficiente,
- (iii) se não há nenhum outro objeto da classe C que chegue a satisfazê-la(s) na mesma ou em maior medida.

Substituindo as variáveis pelos dados e descrições localizadora e caracterizadora de um nome próprio qualquer, nós estabelecemos o que pode ser chamado de a *regra de identificação* (RI) para esse nome. Uma regra que estabelece as condições necessárias e suficientes para a sua aplicação.

Assim, se RMD for aplicada ao nome “Aristóteles” teremos a seguinte regra de identificação para esse nome próprio:

RI-Aristóteles:

O nome próprio “Aristóteles” se aplica a seres humanos em um mundo possível qualquer *see* nesse mundo existiu um ser humano que nasceu em Estagira em 384 a.C., viveu grande parte de sua vida em Atenas e morreu em Chalcis em 322 a.C. e/ou ele foi o autor das grandes ideias contidas no *opus* aristotélico, satisfazendo essa condição (ou essas condições) suficientemente e mais do que qualquer outro ser humano.

Essa regra de identificação para o nome próprio “Aristóteles” é intuitiva. Se a aplicarmos a regra ao caso de um filósofo árabe medieval com o pseudônimo “Aristóteles”, que em um mundo possível *m4* muito próximo ao nosso escreveu grande parte do *opus* aristotélico, veremos que essa pessoa satisfaz unicamente e suficientemente a condição B, mas não a condição A, o que já basta para que ela

satisfaça a regra de identificação para Aristóteles. Contudo, se o mundo possível for m_5 , que difere de m_4 apenas por ter existido nele também outro Aristóteles, nascido em Estagira em 383 a.C., filho do médico Nicômaco, mas falecido jovem, pouco após sua chegada a Atenas, onde foi enviado para aprender filosofia com Platão, não saberíamos mais ao certo o que dizer. Ou seja: se a medida da satisfação dos objetos concorrentes for aproximadamente a mesma pode não haver como decidir, o que significa não termos como aplicar a regra, devendo concluir que Aristóteles não existe, posto que na lógica dos conflitos criteriosais $1 + 1 = 0$.

Casos como esse recordam o paradoxo do navio de Teseu, bem conhecido nos manuais de filosofia. Digamos que esse navio tenha sido batizado com o nome de “Calibdu”. Como é sabido, no curso dos anos Teseu repôs pouco a pouco as partes do seu navio até que, no final, todas elas foram substituídas. Contudo, alguém decidiu então recondicionar as partes antigas e com elas construir outro navio igual ao primeiro. Digamos que então alguém pergunte: “Qual dos dois navios é Calibdu?” O paradoxal aqui é que não sabemos bem o que responder. A primeira vista pode parecer que ambos são o navio de Teseu. Mas isso seria contraditório, pois por definição um termo singular não pode se referir a mais de um objeto. Minha proposta é a de que a questão de saber qual dos navios é Calibdu é indecidível devido a um conflito criterial que se dá entre as duas regras-descrições fundamentais para esse nome. O primeiro navio satisfaz uma regra localizadora, que nos diz que o navio de Teseu é aquele que foi construído em um lugar e tempo específicos, tendo então seguido uma determinada carreira espaço-temporal. A segunda regra, satisfeita pelo segundo navio, é caracterizadora. Ela nos diz que o navio de Teseu é aquele que foi construído com certo material. Não temos, por isso, como decidir.

Nesse ponto alguém poderá, com razão, objetar que a regra caracterizadora é mais complexa. Ela inclui características funcionais e estruturais que foram preservadas em ambos os navios. Como consequência, parece que o primeiro navio deve ser o Calibdu, pois ele satisfaz mais completamente as regras-descrições fundamentais. Contudo, podemos equilibrar essa diferença aumentando a rapidez da substituição das partes velhas pelas novas, de modo a encurtar a carreira espaço-temporal do objeto até que a substituição das peças se complete. Se toda a sequência de substituições de partes

tiver lugar em apenas três meses, teremos dúvidas. E se ela tiver lugar em uma semana ou em um dia? Nesses casos com certeza consideraremos o segundo navio como sendo o de Teseu e não mais o primeiro, dizendo que ele foi primeiro desmontado e depois remontado em outro lugar.

Outra questão diz respeito ao nome próprio em sua expressão fonética e morfológica. Imagine um mundo possível em tudo idêntico ao nosso, com exceção do fato de que a pessoa que nasceu em Estagira, que viveu de 384 a 322 a.C., que escreveu o *opus aristotélico*... se chamava Pitacus. É óbvio que reconheceremos esse Pitacus como o nosso Aristóteles nesse mundo possível. Mas parece que Pitacus não satisfaz RI-Aristóteles. Afinal, ele não se chamava “Aristóteles”. A resposta é que a palavra “Aristóteles” no *definiendum* apenas substitui uma classe de possíveis expressões fonéticas e morfológicas de nomes próprios que satisfazem as condições de localização e caracterização. Como já consideramos, a descrição definida “a pessoa de nome N” é apenas uma descrição definida auxiliar, que pode ser abandonada.

Há nessas respostas um elemento de vaguidade capaz de incomodar alguns. Contudo, a vaguidade, a indeterminação semântica, é um pouco como a indeterminação na física quântica: ela é irreduzível. Ela é um elemento geralmente irreduzível de nossos conceitos, uma vez que reflete a indeterminação de suas próprias referências. Ela em nada atrapalha o funcionamento da linguagem, a menos nos contextos onde não é vista como imprecisão. Se quisermos ter uma teoria dos nomes próprios precisaremos abandonar o “preconceito da pureza cristalina” (Wittgenstein) e adotarmos uma teoria que dá conta da indeterminação semântica.

Falando em preconceito da pureza cristalina, é curioso notar que é possível parafrasear RI-”Aristóteles”, assim como qualquer outra regra de identificação de nome próprio, usando artifícios da teoria russelliana das descrições. Para tal precisamos primeiro transformar as descrições fundamentais em predicados. Assim, se o predicado “...ser humano que nasceu em Estagira em 384 a.C., viveu a maior parte de sua vida ativa em Atenas e morreu em Chalcis em 322 a.C.” for simbolizado por A, o predicado “...autor das grandes ideias do opus aristotélico for simbolizado por B, se o predicado “...gostava de cães” for simbolizado por C, nós poderemos (de modo simplificado) formalizar a sentença “Aristóteles gostava de cães” como:

$$\exists x ((Ax \vee Bx) \& (y) ((Ay \vee By) \rightarrow y = x) \& Cx).$$

Aqui o requisito de existência é o da disjunção das condições fundamentadoras suficientemente satisfeitas, enquanto o requisito de unicidade substitui o requisito de maior satisfação de x . Essa paráfrase salienta os aspectos formais, mas é imperfeita uma vez que substitui a condição de maior satisfação suficiente das condições fundamentais pela condição derivada de que um único indivíduo satisfaz essas condições suficientemente.

Problemas insolúveis da concepção causal-histórica

Podemos agora nos perguntar se o metadescritivismo não deveria incorporar alguma coisa da concepção causal-histórica, transformando-se em um metadescritivismo causal.⁶ Meu ponto de vista é o de que isso é desnecessário. Que geralmente existe uma cadeia causal é um fato que mesmo descritivistas como P. F. Strawson há muito reconheceram. A novidade da concepção causal-histórica é a de que ela ambiciona primariamente explicar a referência dos nomes próprios através dessa cadeia causal. O que eu contesto é precisamente essa ambição, sugerindo que o poder explicativo do recurso à cadeia causal externa, se ela existir, não é capaz de ser primário, mas é derivado da explicação que recorre a descrições expressando regras cognitivas ou pré-cognitivas através das quais o objeto de referência é identificado. Além disso, parece claro que nomes próprios podem ter sentido cognitivo sem que o elemento causal-histórico exista, como o demonstram os muitos exemplos de nomes próprios vazios, como, digamos, “Eldorado”. Esse é um nome próprio ordinário. Os espanhóis ouviram dos índios detalhes sobre uma cidade riquíssima, situada em algum lugar ao leste da cordilheira dos Andes, embora nunca a tenham encontrado. Também acontece de nomes próprios possuírem referência sem que tenha sido formada qualquer cadeia causal. Um caso é o de nomes que foram criados antes do aparecimento de seus

⁶ Em uma primeira versão das idéias aqui desenvolvidas tentei equivocadamente incorporar um elemento causal ao metadescritivismo. Ver C. F. Costa: “A Meta-Descriptivist Theory of Proper Names”.

portadores. “Brasília”, por exemplo, é o nome de uma cidade planejada e nomeada antes de ter sido construída, o que torna impossível que o objeto seja a causa determinante das aplicações iniciais desse nome. Outro caso é o de nomes próprios proferidos sem que haja cadeia causal alguma, mas que foram inferidos. Sabemos, em um exemplo conhecido, que o referente do nome próprio “Ramsés VIII” deve ter existido por volta do século oitavo a.C., pois embora não saibamos nada sobre esse faraó, sabemos algo sobre Ramsés VII e Ramsés IX. Certamente, quando pronuncio o nome “Ramsés VIII” meu proferimento não se encontra no final de uma cadeia causal-histórica...

Um teste para saber se a incorporação de alguma cadeia causal é necessária à explicação da função referencial dos nomes próprios consiste em considerar se há exemplos de proferimentos em que o falante não é bem sucedido em estabelecer nenhuma vinculação causal entre o nome próprio que usa e a sua referência, embora esse nome próprio seja em geral causalmente vinculado a sua referência. Se por causa disso a referência desaparece é sinal de que ela depende necessariamente da associação causal. Imagine, pois, que um psicótico em um sanatório anuncie que os extraterrenos irão pousar na cidade de Saratoga, na Califórnia. Ele associa ao nome “Saratoga” ao menos a descrição “uma cidade na Califórnia”. De fato, existe uma pequena cidade na Califórnia com esse nome. Mas suponhamos que o nome Saratoga tenha sido invocado em sua mente por ter causalmente ouvido por ele em um documentário sobre a batalha de Saratoga na guerra da independência americana, uma batalha que se deu na costa leste dos EUA. Como a palavra “Califórnia” foi pronunciada logo a seguir, ele entendeu que Saratoga é uma cidade situada na Califórnia... Não há, portanto, nenhuma cadeia causal (ao menos em meu exemplo) relacionando a cidade de Saratoga com o nome “Saratoga” pronunciado pela pessoa. A questão é: a pessoa logrou referir-se ao portador do nome? Embora a linguagem natural não nos forneça uma intuição positiva forte para casos inusitados como esse, a resposta é que *podemos* dizer que sim, conquanto acrescentemos que a referência se deu por acaso. Podemos dizer que a pessoa foi bem sucedida em se referir à cidade de Saratoga, embora essa referência tenha sido *meramente coincidental*. Não há, pois, nenhuma intuição linguística que nos obrigue a introduzir um elemento causal no descritivismo, mesmo sendo matéria de fato que as referências sejam em geral causalmente implicadas.

O significado dos nomes próprios

A solução dada ao problema da referência permite-nos responder ao problema do significado dos nomes próprios, entendendo a palavra “significado” no sentido fregeano de “sentido” (*Sinn*) ou “conteúdo informativo” (*informativer Gehalt*) ou ainda (em frases) “valor epistêmico” (*Erkenntniswert*). Em que ele consiste? Segundo a teoria do agregado, o significado do nome próprio residiria nas regras expressas pelas descrições definidas que ele abrevia. Podemos agora qualificar melhor essa tese. Que regras dão significado a um nome próprio dado? Certamente, a regra meta-descritiva (RMD) não é capaz disso, pois ela é a mesma para cada nome próprio, não permitindo distinguir o significado de um nome próprio do significado de outro nome próprio. Também as regras expressas pelas descrições auxiliares não são capazes de dar o significado de um nome próprio, embora se possa dizer delas que formam suas “franjas” de significação. O significado de um nome próprio deve residir então, centralmente, naquilo que o distingue dos outros nomes próprios, ou seja, ele deve consistir de suas regras-descrições fundamentais, localizadora e caracterizadora. Quem realmente sabe o significado do nome próprio é quem, em maior ou menor medida, domina essas regras, sendo essa pessoa o que chamamos de seu *usuário privilegiado*. Outras pessoas, como aquele que sabe apenas que Aristóteles foi um pensador grego (descrição indefinida) ou que acredita que ele foi o descobridor da lei da alavanca (descrição errônea, mas convergente) podem ser capazes de inserir o nome “Aristóteles” corretamente no discurso, sabendo mesmo algo de seu sentido, mas não sabem de maneira suficiente o que esse nome significa, o que deveria ser suficiente para a identificação do seu portador, devendo assumir que usuários privilegiados existem e que estes seriam capazes de completar ou corrigir o pouco que eles sabem.

Essa constatação nos permite admitir que o conhecimento do significado de um nome próprio não precisa ser compartilhado entre cada um dos usuários. Ele pode ser tido apenas pelos usuários privilegiados. Além disso, os usuários privilegiados não precisam individualmente conhecer todo o significado do nome, sendo possível que cada um deles tenha acesso a uma parte diferente do significado. É possível até que parte do conteúdo informativo do nome próprio seja mesmo armazenada fora das

mentes humanas, conquanto ela possa ser utilizada por elas, o que demandará um elemento cognitivo-descritivo prévio. A condição a que queremos chamar atenção é apenas a de que o significado conhecido do nome próprio – a sua regra de identificação – não precisa ser conhecido por cada um de seus usuários. Esse significado é propriedade necessária apenas da comunidade linguística formada pelo conjunto de seus usuários, podendo ser atualizado na soma dos elementos desse conjunto.

Minha sugestão, pois, é que pessoas que só conseguem associar ao nome descrições auxiliares ou genéricas ou mesmo insuficientemente corretas, só são capazes de se referir ao objeto de um modo *dependente* ou *insuficiente*, por se fiarem na existência de usuários privilegiados do nome. Para o nome “Aristóteles” esses usuários são especialistas, conhecedores da história da filosofia ou da cultura. Só eles, em conjunto ou isoladamente, são capazes de se referir a esse filósofo de modo *independente* ou *suficiente*, por associarem seu nome a descrições fundamentais. Há aqui um equivalente ao que Putnam chamou de divisão do trabalho linguístico, só que essa divisão possui caráter potencialmente ou atualmente cognitivo, sendo inteiramente compatível com o descritivismo.

Tendo isso em mente, em situações nas quais os usuários privilegiados do nome próprio desaparecessem e com eles os próprios meios de se obter o conhecimento das descrições fundadoras, o significado do nome próprio também se perderia. Imagine, pois, que após uma catástrofe natural restasse apenas uma comunidade de nativos em algum lugar do mundo, que esses nativos sejam capazes de falar inglês, mas que nada saibam da cultura norte-americana. Digamos que um deles encontre em uma folha de papel a única referência restante a Feynman no mundo inteiro, a frase “Richard Feynman foi um grande conhecedor de Tannu Tuva”. Claro que nós mesmos somos capazes de saber que ao pensar essa descrição ele se refere insuficientemente ao criador da eletrodinâmica quântica (basta digitarmos no Google “Richard Feynman” e “Tannu Tuva”). Mas não é esse o ponto que desejo considerar. O que desejo é apontar para o fato de que esses nativos, em sua sociedade, não serão capazes de fazer nada com a descrição encontrada, posto que as condições últimas de referência se tornaram irrecuperáveis para eles: a comunidade linguística à qual eles pertencem não possui as regras-descrições fundamentais para a identificação do portador do nome Richard

Feynman, não possuindo, portanto, a regra de identificação para esse nome próprio. Como resultado, eles próprios não serão capazes de fazer coisa alguma com esse nome.

Essas sugestões permitem-nos responder aos contraexemplos de Kripke ao descritivismo, segundo os quais podemos nos referir a Feynman através de uma descrição indefinida e a Einstein através de uma descrição errônea. Minha sugestão é que essas pessoas são capazes de fazer uma referência incompleta, um gesto em direção à referência, e que isso muitas vezes é tudo o que precisamos. Mas para que tal aconteça é preciso ao menos duas coisas. Primeiro, é preciso que a descrição que a pessoa associa ao nome próprio seja *convergente*, entendendo por descrição convergente aquela capaz de identificar ao menos a classe C a que pertence o portador do nome. Se alguém crê que Feynman foi um grande cientista, essa descrição indefinida é convergente, pois já contém informação sobre Feynman (um homem, um cientista). Se alguém crê que Einstein foi o inventor da bomba atômica, mesmo que essa descrição seja errônea, ela nem por isso deixa de ser convergente, pois ela já implica que Einstein foi um ser humano e que ele foi um cientista, o que é verdadeiro. O mesmo não ocorreria se as descrições forem divergentes, por exemplo, se alguém acreditar que Feynman é o nome de uma marca de perfume ou que Einstein é o nome de uma pedra preciosa. Aqui os portadores dos nomes não pertencem à classe identificada.

A segunda condição é a de que a pessoa possua conhecimento tácito do mecanismo de referência dos nomes próprios, da regra meta-descritiva. Com isso ela sabe ao menos que aquilo que sabe das descrições fundamentais exigidas é insuficiente. Com essas duas condições satisfeitas, com o pouco saber convergente que lhe está disponível, com a consciência que ela tem de sua própria falta de conhecimento, a pessoa já será capaz de inserir o nome próprio adequadamente no discurso, em contextos que reconhece como sendo suficientemente vagos, como tantas vezes acontece. Essa é a razão pela qual alguém pode, em um certo sentido (insuficiente) da palavra, se referir a Feynman sabendo apenas que ele foi um grande cientista e a Einstein acreditando que ele foi o inventor da bomba atômica. Na verdade ele está inserindo o nome corretamente no discurso de modo a fazer uma referência por si mesma insuficiente, adventícia, posto que em última análise dependente da comunidade linguística, a qual possui recursos para completar a referência.

Porque nomes próprios são designadores rígidos

A solução sugerida também permite responder à objeção de que a teoria do agregado não dá conta da propriedade do nome próprio de ser um *designador rígido*, que é mais intuitivamente definida como sendo a propriedade que ele tem de *se aplicar a um mesmo objeto em qualquer mundo possível no qual esse objeto exista*.⁷ Para encontrarmos a resposta basta considerarmos atentamente as regras de identificação dos nomes próprios resultantes da aplicação de RMD. No caso do nome próprio “Aristóteles” a regra estabelece uma identidade entre o nome próprio e uma descrição complexa, que é a seguinte:

o ser humano que satisfaz suficientemente e mais do que qualquer outro as condições de ter nascido em Estagira em 384 a.C., ter vivido a maior parte de sua vida em Atenas... e/ou de ter sido o autor das grandes ideias do opus aristotélico.

A identidade é analítica ou necessária, valendo para todos os mundos possíveis. A descrição aqui apresentada, que exprime a regra de identificação para Aristóteles, é por sua vez um designador rígido: “o ser humano que satisfaz suficientemente e mais do que... o autor das grandes ideias do opus aristotélico” aplica-se em todos os mundos possíveis nos quais Aristóteles existe, pois não há como conceber a sua existência sem conceber a aplicabilidade dessa regra.

Certamente, haverá mundos possíveis nos quais não saberemos se a regra de identificação para um nome próprio é minimamente satisfeita ou não (digamos que em um desses mundos, na corte de Felipe em Estagira em 384 a.C., tenha nascido um feto anencefálico chamado Aristóteles, filho do médico da corte, mas que ele tenha morrido dias após o nascimento... e que ninguém jamais tenha escrito o *opus* aristotélico). Mas isso sugere apenas que a semântica dos mundos possíveis deve ser reescrita de modo a dar lugar a casos indecidíveis. Para dar conta disso o designador rígido precisa ser

⁷ O próprio Kripke oscila entre essa definição e a definição segundo a qual o designador rígido é o que se aplica em todos os mundos possíveis, inclusive naqueles nos quais o objeto não existe. Mas parece claro que no último caso a maioria dos nomes próprios deixaria de ser designadores rígidos, posto que nenhum nome próprio irá referir em um mundo possível no qual o seu objeto de aplicação não existe.

redefinido como sendo aquele que se aplica a todos os mundos possíveis nos quais o objeto *definidamente* existe.

Com isso fica fácil explicar porque nomes próprios são designadores rígidos do ponto de vista do descritivismo. É que as regras de identificação dos nomes próprios, quando expressas por descrições, originam descrições rígidas, aplicáveis em todos os mundos possíveis nos quais esse objeto definidamente existe. A descrição definida acima exprime também um *critério definitório* de aplicação do nome próprio, ou seja, a condição necessária e suficiente para a sua aplicação. Dizer que Aristóteles definidamente *existe* em um mundo possível é o mesmo que dizer que a regra de identificação para esse nome próprio é efetivamente aplicável nesse mundo, pois é a aplicabilidade da regra semântico-conceitual de um nome próprio em um mundo possível aquilo mesmo que entendemos pela existência do seu objeto de referência nesse mundo.

Porque descrições definidas são designadores flácidos

A introdução de regras de identificação para nomes próprios como resultado da aplicação de RMD nos permite explicar não só porque nomes próprios são designadores rígidos, mas porque as descrições definidas que a eles associamos são em geral designadores flácidos, ou seja, designadores que se referem a objetos diferentes em diferentes mundos possíveis. Considere a descrição definida “o fundador do Liceu”. Podemos conceber um mundo possível no qual o nome “Aristóteles” se aplica a Aristóteles, que nele existiu, mas no qual a descrição “o fundador do Liceu” se aplica a outra pessoa, digamos, ao seu discípulo Teofrasto, que nesse mundo foi quem realmente fundou o Liceu. E também podemos conceber um mundo possível em que o nome “Aristóteles” se aplica a Aristóteles, que nele existiu, mas no qual a descrição “o fundador do Liceu” não se aplica, pois nesse mundo nenhum Liceu foi fundado. Por que é assim?

A resposta encontra-se à mão. Sabemos que nenhuma das descrições do agregado, mesmo as descrições fundamentais, se encontra necessariamente vinculada à aplicação do nome próprio, caso o objeto a ser referido pelo nome próprio exista. O que

necessariamente se encontra vinculado à aplicação do nome próprio, caso a sua referência exista, é apenas a regra de identificação do nome próprio, que como um todo estabelece o que chamo, seguindo Wittgenstein, de o *critério definitório* para a sua aplicação. O critério definitório ou primário é aquele que, uma vez dado, garante a existência daquilo de que é critério. Ele se distingue do *sintoma* ou critério secundário, que uma vez dado apenas torna provável a existência daquilo de que é critério.⁸ O contraste entre a descrição que exprime a regra de identificação do nome próprio e as descrições constitutivas do agregado por ele abreviado é feito pela distinção entre critério definitório ou primário e sintoma ou critério secundário. Sempre que a descrição da regra de identificação é satisfeita, sempre que o objeto existe, por nos prover de um critério definitório a regra garante a aplicação do nome próprio a esse objeto. Ora, isso o torna o nome próprio um designador rígido, posto que aplicável em qualquer mundo possível no qual seu objeto de referência exista. Mas quanto a qualquer das descrições do agregado, mesmo as descrições fundamentais (tomadas isoladamente), a sua aplicação consiste na satisfação de sintomas, de critérios secundários meramente probabilizadores da presença do objeto. Assim, a satisfação das descrições definidas associadas ao nome próprio, sejam elas quais forem, apenas probabiliza a aplicação do nome próprio, o que as impede de serem aplicadas em todos os mundos possíveis nos quais o objeto referido pelo nome existe. O resultado disso é que as descrições do agregado são flácidas, posto que podem se aplicar a outros objetos, que não o referente do nome próprio em outros mundos possíveis, ou simplesmente a nenhum objeto, mesmo em um mundo possível no qual o objeto referido pelo nome próprio exista e vice-versa. Aliás, as descrições do agregado não precisam se aplicar não só em outros mundos possíveis (em situações contra-factuais), mas até mesmo ao nosso próprio mundo, uma vez que podemos estar errados em nosso conhecimento dos dados a ele concernentes. Pode ser descoberto, por exemplo, que Aristóteles não foi realmente amante de Herphylis.

⁸ Essa é uma maneira de interpretar o que Wittgenstein escreve sobre a distinção entre critério e sintoma. Ver L. Wittgenstein, *The Blue and the Brown Books*, p. 24 ss.

A evidência a favor da explicação acima apresentada é que a oposição nome próprio rígido *vs.* descrição definida flácida se mantém apenas aonde as descrições vem associadas a nomes próprios e as suas aplicabilidades podem contrastar. Isso nos permite prever que descrições definidas que não se encontram associadas a nomes próprios, especialmente quando são fundamentais, devem funcionar como designadores rígidos. Considere, por exemplo, a descrição “o terceiro regimento de cavalaria de Sintra”. Ela exprime as regras de localização e caracterização do regimento, que por sua vez não possui um nome próprio. Por ser assim ela se nos apresenta como uma descrição definida rígida, aplicando-se em qualquer mundo possível no qual esse regimento exista, mesmo que composto por diferentes cavaleiros e cavalos. Ela é rígida porque exprime uma regra de identificação que não está associada a nenhum nome próprio, não podendo por isso em nenhum mundo possível haver divergência entre o objeto de aplicação dessa regra e o objeto de aplicação de um nome próprio ao qual ela esteja associada. Outros exemplos de descrições naturalmente rígidas são: “o assassinato do arquiduque Ferdinand em Sarajevo em 1914”, “o ponto mais oriental da América Latina” e “a última idade do gelo”, que designam respectivamente um evento, um local e um processo. Elas se aplicarão em qualquer mundo possível no qual alguém assassinou o arquiduque em 1914 (mesmo que o arquiduque seja outro e que o assassino não seja Graviolo Princeps), em qualquer mundo possível no qual exista alguma América Latina com algum ponto mais oriental (mesmo que ele esteja na Terra do Fogo e não em João Pessoa), e em qualquer mundo no qual tenha havido uma última Idade do Gelo (mesmo que ela tenha terminado há 60.000 anos e não há 10.000 anos).

Respostas a contraexemplos

A teoria meta-descritivista dos nomes próprios recém esboçada permite respostas mais convincentes aos contraexemplos levantados contra o descritivismo. Responderei aqui apenas aos mais influentes.

Consideremos primeiro o caso inicialmente mencionado de nomes semificcionalis, como Robin Hood. Kripke sugeriu que esses nomes demonstram a verdade da concepção causal-histórica, pois embora não tenhamos descrições definidas

capazes de identificar seus portadores, sabemos que eles se referem a algum objeto que foi a fonte causal desses nomes.

Nossa resposta começa com a constatação de que nomes semi-ficcionais abreviam dois tipos de descrições: as não-ficcionais, que servem para identificar o objeto que realmente originou causalmente o uso do nome; as ficcionais, que foram adições imaginativas posteriores, nada tendo a ver com o objeto originador do nome. Embora alguma coisa das descrições não-ficcionais seja sabida (sabemos que Robin Hood teria sido um justiceiro que viveu na Inglaterra no período medieval) em geral não sabemos distinguir quais são as descrições ficcionais e quais são as não-ficcionais, nem a extensão disso.

Imagine agora que uma das muitas teorias já propostas sobre quem teria sido Robin Hood seja demonstrada. Suponhamos que documentos sejam descobertos comprovando a teoria de L. V. D. Owen, de acordo com a qual o Robin Hood histórico foi um fora da lei chamado Hobbhod que viveu na primeira metade do século XI em Yorkshire. Nesse caso ao menos nossa descrição caracterizadora de Robin Hood como um fora da lei seria confirmada e complementada, enquanto nossa descrição localizadora de Robin Hood como tendo vivido no período medieval na Inglaterra teria sido confirmada e precisada. O nome próprio teria a sua referência real comprovada com base em descrições e não no elemento causal, mesmo que este também exista.

Suponhamos agora, para prover um contraexemplo à explicação causal-histórica, que se descubra que o nome Robin Hood não tenha sido invocado por um ser humano na mente do primeiro escritor medieval a usar o nome, mas por outro ser vivo, qual seja, pelo seu bravo cão perdigueiro de nome Robin, que costumava acompanhá-lo em suas incursões na floresta de Sherwood... Nesse caso não diremos que o nome próprio Robin Hood se refere ao cão de caça do escritor, mas que esse nome não possui referência alguma, pois ele é na verdade um nome completamente ficcional. Nossa teoria oferece uma explicação para o que acontece: não se trata somente do fato de que o cão não satisfaz coisa alguma da regra de caracterização que temos para Robin Hood, mas, essencialmente do fato de que nossa regra de identificação para Robin Hood se aplica à classe C dos seres humanos e não à classe dos canídeos. Se considerarmos agora a teoria causal-histórica, parece que por meio dela devemos reconhecer o nome “Robin Hood”

como permanecendo semificcional por se referir ao cão do escritor medieval. Afinal, há uma cadeia causal-histórica que começou com o cão batizado com o nome de Robin... Essa é, porém, uma conclusão absurda.⁹

Outro famoso contraexemplo de Kripke é o de um falante que associa ao nome do matemático Kurt Gödel a descrição “o inventor da prova da incompletude”. Imagine, escreve ele, que se descubra que essa prova foi na verdade descoberta por Schmidt, que morreu em Viena em circunstâncias misteriosas, e que o seu amigo Gödel tenha roubado a prova e publicado em seu próprio nome. Nesse caso, se nomes fossem abreviações de descrições, pensa Kripke, uma vez informada disso a pessoa deveria admitir que Gödel é Schmidt, pois é a Schmidt que devemos agora associar a descrição. Mas isso é contra-intuitivo, pois a pessoa continuará certa de que Gödel é Gödel e não Schmidt, mesmo sabendo que Gödel foi um falsário que não descobriu nenhuma prova da incompletude.

10

A resposta que a versão meta-descritivista da teoria do agregado dá ao exemplo em questão é, diversamente do esperado, perfeitamente intuitiva. O usuário privilegiado do nome “Gödel” o reconhece por satisfazer a regra de localização (A) de ter nascido em Brunn em 1906, estudado em Viena, emigrado para os EUA e trabalhado em Princeton, onde faleceu em 1978, e por satisfazer a regra de caracterização (B) de ter sido um grande matemático que descobriu o teorema da incompletude, além de ter feito muitas outras contribuições menores. Assim, mesmo que Gödel deixe de satisfazer parte (digamos 2/3) da regra de caracterização, ele continua satisfazendo integralmente a regra de localização, satisfazendo, pois, RI para “Gödel” bem mais do que RI para “Schmidt”. Eis porque Gödel não pode ser Schmidt!

⁹ Imagine que o ser vivo que acompanhasse o escritor medieval à caça fosse outra pessoa, um amigo chamado Hode, que lhe inspirou a estória de Robin Hood. Pela concepção causal, a pessoa batizada com o nome “Hode” seria o verdadeiro Robin Hood, pois é a causa do uso do nome semificcional, o que é escassamente intuitivo. Pela teoria proposta, porém, a resposta irá depender de detalhes adicionais, no caso, se Hode satisfaz suficientemente a condição de predominância e se ele satisfaz a condição de predominância em contraste com o objeto meramente imaginado pelo escritor medieval. Assim, se Hode realizou certos atos de bravura e compaixão, parece que há razões para concluir que Hode foi o Robin Hood originário.

¹⁰ Saul Kripke: *Naming and Necessity*, pp. 83-84.

Quanto à pessoa que associa ao nome “Gödel” somente a descrição “o inventor da prova da incompletude”, como falante competente ela tacitamente conhece a forma da regra de identificação para nomes próprios e, por conhecê-la, ela sabe que a regra que essa descrição exprime é muito incompleta, apenas parte da regra caracterizadora, recusando-se por isso a aceitar que Gödel é Schmidt enquanto não obtiver maiores informações.

Uma curiosidade acerca do exemplo é que, como ao menos parte de uma das duas descrições fundamentais identificadoras de Gödel é satisfeita por Schmidt, é possível dizer que este último passa a herdar alguma coisa do significado do nome “Gödel”, mesmo que não ganhe a sua referência. E isso realmente acontece. Digamos que um lógico, revoltado pela notícia acerca do roubo do teorema e com pena de Schmidt, lance a exclamação “Schmidt é quem foi o verdadeiro Gödel!” Essa é uma frase verdadeira se for entendida como uma hipérbole. E a razão pela qual ela é verdadeira é dada por nossa versão da teoria descritivista, a qual prevê que o nome Schmidt herda alguma coisa relevante do significado do nome “Gödel”.

Há, por fim, uma maneira de fazer com que Gödel seja realmente Schmidt, muito embora ela dê a Kripke o bolo sem o direito de comê-lo. Imagine que bem no início da estória Schmidt, por alguma razão, tivesse assassinado o jovem Gödel e assumido a sua identidade. Schmidt, que era muito melhor lógico que Gödel, descobriu então incompletude da aritmética, casou-se com Adele, fugiu para os EUA pela Transiberiana em 1940, tornou-se professor em Princeton e faleceu em 1978, de modo que aquele sujeito de calças curtas junto a Einstein na famosa foto de ambos era ele mesmo, o falsário Schmidt. Nesse caso não há dúvida de que Gödel é Schmidt. E o metadescritivismo explica: ele é Schmidt porque as regras-descrições caracterizadora e localizadora, com exceção das descrições relativas à infância, são as de Schmidt e não as da criança que uma vez foi chamada de Gödel, a qual há muito deixou de existir.

Uma última objeção de Kripke que quero aqui considerar consiste na sugestão de que o nome próprio for definido como abreviação de uma descrição definida e formos tentar explicá-la, nós teremos de recorrer outra vez ao nome próprio, caindo em um círculo vicioso. Assim, se associo o nome próprio “Einstein” à descrição “o criador

da teoria da relatividade”, ao dizer o que é a teoria da relatividade terei de dizer que ela foi a teoria criada por Einstein.

À luz de um descritivismo mais complexo essa objeção perde seu poder de convicção. A descrição “o autor da teoria da relatividade” contém o principal da descrição caracterizadora, da razão pela qual nos lembramos de Einstein. Mas é perfeitamente possível explicar a teoria da relatividade sem mencionar o nome “Einstein”. O mesmo também acontece com a descrição localizadora para esse nome: “a pessoa nascida em 1879 em Ulm, que estudou em Zurique, lecionou em Berlin e, a partir de 1933 se tornou professor em Princeton, onde faleceu em 1953”. Nenhuma circularidade. Certamente, nosso aprendizado do sentido, do conteúdo informativo de um nome próprio não é imediato, mas progressivo, razão pela qual recorreremos outras vezes ao nome próprio sempre que pedimos um detalhamento da explicação. É a esse “mistério” da recursividade do *definiendum* no aprofundamento do *definiens* que apela a objeção de Kripke.

Quero, por fim, analisar rapidamente um contra-exemplo proposto por Keith Donnellan.¹¹ Imagine, escreve ele, que se descubra que Tales não foi na verdade nenhum filósofo, mas um sábio cavador de poços, que cansado de sua profissão uma vez dissera: “Quem me dera se tudo fosse água para eu não ter de cavar esses malditos poços”, tendo essa frase passado equivocadamente a Heródoto, a Aristóteles e a outros como veículo da ideia atribuída ao filósofo Tales, segundo a qual a água é o princípio de tudo. Digamos também que a ideia de que tudo é água tenha sido sustentada por um eremita que viveu tão remotamente que nem ele nem suas doutrinas tenham qualquer conexão histórica conosco. Mesmo assim nós não diremos que Tales foi o eremita. A teoria causal-histórica possui uma maneira de explicar isso. Segundo ela é assim porque foi o cavador de poços Tales quem se encontrava no princípio da cadeia causal-histórica e não o eremita. Mas segundo a teoria descritivista o eremita é quem deveria ser Tales, pois é ele quem satisfaz a descrição – um resultado contraintuitivo.

A resposta que a teoria metadescritivista aqui proposta irá oferecer parte da constatação de que em certos casos a descrição da história causal simplesmente faz

¹¹ Keith S. Donnellan: “Proper Names and Identifying Descriptions”, pp. 373-375.

parte da descrição caracterizadora. Esse é precisamente o caso de Tales, pois o que mais nos importa na formação da regra caracterizadora para Tales é o seu lugar e influência na origem da filosofia ocidental. A descrição caracterizadora de Tales não poderia se resumir à ridícula afirmação de que tudo é água *per se*, pois se um filósofo de uma época mais próxima à nossa escrevesse isso ele seria visto como incompetente. A descrição caracterizadora de Tales só importa para nós por incluir a história causal. Podemos resumi-la como: “a pessoa *que originou a doxografia encontrada em Heródoto, Aristóteles e outros*, que a descreve como o primeiro filósofo grego, a defender que a água é o princípio de todas as coisas, que tudo é vivo, que tudo é um etc.” Quanto à regra localizadora, sabemos que Tales foi “o milesiano que viveu provavelmente de 624 a 547-8 a.C...” Em vista disso, se retornarmos ao exemplo de Donnellan, concluiremos que o eremita não pode ter sido Tales. Primeiro porque não satisfaz a descrição localizadora. Depois porque ele não satisfaz a descrição caracterizadora, mesmo que satisfaça alguma coisa dela. Assim, mesmo que Tales tenha sido um cavador de poços milesiano que viveu de 624 a 547-8 a.C., ele satisfaz as regras fundamentadoras muito mais completamente do que o eremita. Afora isso é preciso notar que dependendo dos detalhes que forem adicionados ou subtraídos ao exemplo dado, as nossas intuições podem se alterar, levando-nos tanto à conclusão de que nenhum Tales realmente existiu quanto, eventualmente, à conclusão de que Tales na verdade foi o eremita.

Acredito que para quem não quiser se curvar ao argumento de autoridade que secretamente rege a discussão filosófica, para quem for capaz de pensar por si mesmo a despeito da imensa influência imposta pela nova ortodoxia causal-externalista já de há muito vigente, não será difícil admitir que a teoria dos nomes próprios aqui esboçada é a mais auspiciosa.

Referências

Costa, C. F.: “A Meta-Descriptivist Theory of Proper Names”, *Ratio (New Series)* XXIV, 3, 2011, 259-281.

- Devitt, M.: “Meanings just ain’t in the Head”, in George Boolos: *Meanings and Method: Essays in Honor of Hilary Putnam* (Cambridge: Cambridge University Press 1990).
- Devitt, M.: *Designation* (New York: Columbia University Press 1981).
- Donnellan, K. S.: “Proper Names and Identifying Descriptions”, in D. Davidson e Gilbert Harman: (Dordrecht/Boston: Reidel 1972).
- Dummett, M.: *Frege: Philosophy of Language* (London: Duckworth 1981, 2. ed.).
- Frege, G.: “Der Gedanke”, *Beiträge zur Philosophie des deutschen Idealismus* I, 2, 1918, 58-77.
- Frege, G.: “Sinn und Bedeutung”, *Zeitschrift für Philosophie und philosophische Kritik*, NF 100, 1892, 25-50.
- Haack, S.: *Philosophy of Logics* (Cambridge: Cambridge University Press 1978).
- Jackson, F.: “Reference and Description Revisited”, in J.E. Tomberlin (ed.): *Philosophical Perspectives 12: Language, Mind and Ontology* (Oxford: Blackwell 1998), 201-18.
- Kripke, S.: *Naming and Necessity* (Cambridge MA: Harvard University Press 1980).
- Lewis, D.: “Putnam’s Paradox”, *Australasian Journal of Philosophy* 62, 1984, 621-36.
- Russell, B.: *The Problems of Philosophy* (Oxford: Oxford University Press 1980 (1912)).
- Russell, B.: “On Denoting”, *Mind* 14, 1905, 479-493.
- Searle, J. R.: “Proper Names”, *Mind* 67, 1958, 166-173.
- Searle, J. R.: “Proper Names and Descriptions”, Paul Edwards (ed.): *The Encyclopedia of Philosophy* (New York/London: Collier MacMillan 1967), vol. 6, 487-491.
- Searle, J. R.: *Intentionality: an Essay in the Philosophy of Mind* (Cambridge: Cambridge University Press 1983).
- Strawson, P. F.: *Individuals: An Essay on Descriptive Metaphysics* (London: Routledge 2003(1959)).
- Tugendhat, E.: *Vorlesungen zur Einführung in die sprachanalytische Philosophie* (Frankfurt: Suhrkamp 1976).
- Tugendhat, E.: *Propedêutica Lógico-Semântica* (Petrópolis: trad. bras. Vozes 1997).
- Wittgenstein, L.: *The Blue and the Brown Books* (Oxford: Basil Blackwell 1958).

Investigação Filosófica: vol. 4, n. 1, artigo digital 5, 2013.

Wittgenstein, L.: *Philosophische Untersuchungen* (Frankfurt: Suhrkamp 1983).

Uma analítica do poder – Conversa com Judith Butler

Por *Claire Pagès & Mathieu Trachman* [04-12-2012]

Tradutores: Cristiane Maria Marinho, Alba Liarth da Cruz,
Kácia Natalia de Barros Sousa Lima & Leonardo Lima Ribeiro

Por ocasião da publicação de seu último livro, *Parting Ways. Jewishness and the Critique of Zionism* [Judaísmo e a crítica do Sionismo], Judith Butler retoma nesta entrevista questões sobre seus mais recentes trabalhos: criticar a precariedade crescente que implicam as evoluções recentes do capitalismo; defender a possibilidade de uma vida radicalmente democrática, atenta à multiplicidade das relações de poder.

Desde seus trabalhos sobre gênero, no início dos anos 1990, os campos de pesquisa de Judith Butler se diversificaram consideravelmente. A análise dos modos de subjetivação permanece como um dos fios condutores de seus trabalhos, mas a filósofa tem ido além da teoria feminista para tratar de questões éticas e religiosas. Em *O Estado Global [L'État global]*, Judith Butler apresenta seu projeto como um “nova analítica do poder” (p.93). Isso implica repensar tanto a questão da territorialidade quanto da soberania, escutar os diferentes destinos da soberania para, a partir deles apresentar hoje “o mapa emergente”. Abordando questões de Estado, do neoliberalismo, do lugar da religião no espaço público, ela se pergunta sobre o que torna uma vida vivível. Se os textos da tradição filosófica são mobilizados, suas análises estão igualmente ancoradas numa atualidade política que exige por vezes a tomada de posição. É o caso, por exemplo, do casamento gay ou do conflito israelita-palestino.

Precariedade sociológica, precariedade ontológica

La Vie des Idées [A vida das Ideias]: Em seus trabalhos, particularmente em *O que faz uma vida*, você propõe uma análise aprofundada da ideia de precariedade. Ela é bastante real, mas é menos, ao que parece, uma noção sociológica do que ontológica. Nosso questionamento, então, seria este: o neoliberalismo e a crise econômica de grande parte oriundos da desregulamentação neoliberal dos mercados não estão na origem de uma nova forma de precariedade? Ou então se trata do fator conjuntural que faz inverter a precariedade constitutiva do ser humano em uma precariedade insuportável? A incidência específica do econômico nos interessa tanto que você insiste, em *O Estado Global*, sobre a necessidade de não eludir ou negligenciar a economia em proveito de uma análise estrita das causas políticas.

Judith Butler: Concordo com a formulação que você propõe, isto é, que o fator econômico transforma a precariedade constitutiva do ser humano em uma precariedade insuportável. Minha única incerteza recai sobre a questão de saber se, para retomar a sua terminologia, o “ontológico” é separável do “sociológico”? Temos aqui, portanto, dois problemas, pois se afirmamos que o neoliberalismo tende a tornar as pessoas descartáveis e expõe as populações à precariedade, devemos nos perguntar se designamos por “neoliberalismo” uma lógica e um sistema de poder puramente econômicos ou ainda um regime de poder que rege as práticas de formação do sujeito, inclusive de si próprio, assim como o fato de que a valorização do parâmetro de instrumentalidade integra e ultrapassa, doravante, a esfera convencional do “econômico”. O poder e a onipresença do “neoliberalismo” nos forçam a pensar a heteronomia do econômico e a forma com a qual as lógicas que governam as suas operações ultrapassam a esfera puramente econômica? Devemos renunciar a idéia de uma esfera puramente econômica por causa do neoliberalismo, ao mesmo tempo em que não podemos nos abster do econômico?

É mais difícil responder a sua pergunta sobre a ontologia, mas quero novamente insistir no fato de que uma reflexão sobre a precariedade visa antes de tudo apreender mais fundamentalmente o que significa ser um ser social. Visto que a sociologia se apoia sempre nesse postulado, importa pensá-lo de forma crítica: o que queremos dizer quando nos referimos ao conceito sociológico de sujeito, e como o distinguimos tradicionalmente do conceito ontológico de sujeito? Quando eu defendo que o sujeito é constituído socialmente, ou que o sujeito é constituído em e por suas relações sociais com os outros, é uma afirmação sociológica ou ontológica? Para mim, a ontologia não se situa em outro nível senão o social nesse tipo de debate, porque eu tento dizer que as “criaturas” humanas – e não somente os humanos – dependem fundamentalmente das instituições sociais para sua sobrevivência e subsistência. Isso significa que quaisquer que “sejam” esses seres, o que eles “são” é constituído no cruzamento dessas relações, o que implica também que, quando as instituições sociais falham, eles ficam ameaçadas de “não-ser” ou de formas de morte social. Poderíamos chamar esse processo de uma ontologia social, mas as formas de dependência e vulnerabilidade a respeito das instituições sociais têm tendência a variar, e uma análise que se ativesse a um “único nível de ontologia” não seria possível.

La Vie des Idées: Em *O que faz uma vida [Ce qui fait une vie]*, você investiga as condições concretas, sociais e políticas de uma “vida vivível”. Questão inicial da satisfação das necessidades elementares e da necessidade de proteção, assim como estar inserido nas relações humanas de sociabilidade. Você menciona igualmente a necessidade de estar inserido em uma rede de trabalho. Nós gostaríamos de saber mais sobre o lugar e a função do trabalho em sua reflexão. Trata-se de uma condição essencial da “vida vivível” e de uma estrutura de reconhecimento fundamental para o sujeito? Ou melhor, o trabalho é uma condição mais contingente, uma necessidade factual para o indivíduo e não uma estrutura simbólica de sua existência?

Judith Butler: É evidente que o trabalho é necessário para a reprodução da pessoa – tomo emprestada essa ideia da teoria da produção em *A ideologia alemã* de Marx – e essa ideia permanece central em minha reflexão. O trabalho é também necessário para produzir as condições materiais da existência e da subsistência. Assim, sou, por exemplo, a favor de um “direito do trabalho” e creio que os Estados têm a obrigação pública de criar condições que permitam às populações trabalhar, caso elas possam. Oponho-me às formas de capitalismo protestante que afirmam que apenas aqueles que trabalham merecem ser alimentados e abrigados, pois sustento que esses direitos fundamentais deveriam ser garantidos pelos Estados, independente do fato dos indivíduos terem ou não um trabalho. Assim, recuso dizer que apenas o trabalho fornece as condições materiais da produção da vida humana. Isso seria defender uma posição moral em contradição com a obrigação pública de fornecer, entre outras coisas, moradia, alimentação, cuidados médicos, educação.

Quando falamos de formas induzidas de precariedade, designamos formas de organização do trabalho que se instalam no emprego aleatório e no caráter substituível e descartável da mão de obra. Essas formas de precariedade são produzidas e calculadas para fornecer uma força de trabalho “flexível”, e elas induzem, em uma vasta escala, a insegurança e o desespero entre os trabalhadores. É também uma forma de impedir toda projeção em direção ao futuro e de produzir uma estrutura de dívida permanente para aqueles que não têm meios de prever qual será seu trabalho no futuro.

O trabalho das normas

La Vie des Idées: Você trabalhou muito para descobrir o que age como uma norma, em primeira instância no gênero, sem que isso fosse dito. Em uma linha foucaultiana, você procurou desvelar, do interior, os fenômenos de “literalização” ou de naturalização das normas. Hoje, no quadro da analítica do poder que você

elabora, você se atém a definir uma perspectiva normativa, engajando-se em direção a uma postura da crítica social. Tal postura implica uma posição de exterioridade, a posição de uma norma em relação a qual é avaliado um estado da realidade – posição da qual você freqüentemente demonstrou o caráter problemático. O que lhe conduziu a insistir sobre a dimensão normativa de sua reflexão e como você concebe a mesma?

Judith Butler: Foucault sempre fez parte de minha reflexão, e esse é o caso ainda hoje. Mas não sigo Foucault tal como se seguisse um pensamento religioso. Adapto sua obra extraordinária a novos fins; e é ele, aliás, sem dúvida, um daqueles que me mostrou que era possível fazer esse uso de outros pensadores. Em todos os casos, a análise da performatividade do gênero sempre se esforçou em mostrar que considerávamos algumas performances como “reais” e outras como “irreais”. Tomei posição contra esta concepção da produção do gênero e adiantei que as apresentações de gênero mais “normativas” e as mais “convincentes” estavam fundadas na mesma lógica mimética daquelas que considerávamos de maneira convencional como desviantes e inacreditáveis. Assim, a ideia do “normativo” intervém duas vezes; no primeiro caso, como você o sugere, a normatividade, como heteronormatividade, designa um processo de normalização e literalização. Mas, no segundo caso, temos um quadro normativo que busca contestar e substituir a distinção propriamente dita do real e irreal. Igualmente quando falo das vidas que se pode chorar e das que não se pode chorar. Isso constitui o elo entre meus trabalhos sobre a política LGBTQ e meus trabalhos mais recentes sobre a guerra. Minha opinião é a de que tem sido um erro considerar que algumas vidas são mais reais, mais vivas que outras, que seriam menos reais, menos vivas. É uma forma de descrever e de avaliar a distribuição diferencial da “realidade” em função do nível de conformidade dessas populações referente às normas estabelecidas. É também uma tentativa de produzir novos esquemas normativos que implicam uma crítica rigorosa da misoginia, da homofobia, do racismo para fazer emergir um mundo

social e político que se caracterizaria pela interdependência, igualdade e mesmo pela democracia radical.

La Vie des Idées: Recentemente, nos Estados Unidos, Barack Obama se pronunciou em favor do casamento homossexual. Na França, o novo presidente François Hollande se disse favorável à abertura do direito ao casamento e à adoção pelos casais homossexuais. Esse reconhecimento político dos casais do mesmo sexo tem sido discutido há muito tempo, alguns vêem nisso uma normalização da homossexualidade. Essas proposições aparecem igualmente em um momento em que o homonacionalismo – a instrumentalização das questões sexuais nas relações de raça e a retórica dos conflitos das civilizações – torna-se uma questão crucial da política sexual. Como você analisa essas tomadas de posição dos governos de esquerda nos Estados Unidos e na Europa? Quais são as questões políticas envolvidas nisso?

Judith Butler: Nos Estados Unidos, a posição a favor do casamento gay tem tido a tendência de instalar uma nova normatividade no seio da vida gay, resultando em recompensas para gays e lésbicas que adotem a vida a dois, a propriedade e as liberdades burguesas, o reconhecimento público. É necessário ser a favor do casamento gay, e eu sou. Mas o que me preocupa é o fato de que esse assunto tenha se tornado mais importante do que outros objetivos políticos, em particular o direito dos transexuais em se resguardar da violência, inclusive da violência policial, da formação continuada, de ação social e do tratamento do HIV, a necessidade de serviços sociais para pessoas LGBTQ que não são casadas, uma política sexual radical que não se calque nas normas matrimoniais predominantes. É seguramente algo bom que os gays e lésbicas tenham esse direito, caso eles escolham exercê-lo. E de forma totalmente distinta, eu sou decididamente a favor do direito de toda pessoa em ter acesso à adoção e às tecnologias reprodutivas,

independentemente do estado civil ou de orientação sexual. Trata-se de formas fundamentais de oposição à discriminação e eu defendo isso.

É verdade que alguns governos reivindicam seus próprios direitos, estrategicamente, no momento em que negam os direitos aos imigrantes, ou então que travam uma guerra material e cultural contra as populações mulçumanas. Isso suscitou certo número de debates sobre o marketing das empresas em favor da causa gay (*pinkwashing*) e sobre o homonacionalismo. É preciso assegurar que a luta em favor de uma série de direitos minoritários não sirva para privar os direitos de outra minoria. Isso significa que, tão necessária quanto seja nossa luta pelos direitos LGBTQ, ela deva também se inserir no contexto de uma luta pela justiça social e econômica. É necessário, portanto, nos perguntarmos se nossas reivindicações políticas não podem ser utilizadas contrariamente aos engajamentos maiores como os nossos relacionados à solidariedade e à justiça.

Uma política de interdependências

La Vie des Idées: Em *O Estado Global*, para colocar o problema da articulação entre vida e política, você explica que é necessário pensar formas outras de pertencimento que não aquelas da nação e do Estado. Discutindo Agamben, você destaca especialmente que seus trabalhos não permitem compreender as subjetividades apátridas ou militantes. A nova analítica do poder exige finalmente se repensar a noção de soberania, de propor um “novo mapa da soberania”. De qual maneira o conceito tradicional de soberania deve, no seu ponto de vista, ser transformado?

Judith Butler: Eu não sou verdadeiramente uma teórica da soberania, e não estou, portanto, segura de poder responder bem a sua pergunta. A política é um campo complexo e me apoio em vários pensadores para conceber com justeza as noções que não fazem parte diretamente do meu campo de visão. Isso é, certamente, um

limite, mas suponho que todos tenham tais limites. Creio que no contexto que você evoca, tentei dizer que aqueles que são apátridas, que vivem nos campos fronteiriços ou sob uma ocupação, adquirem uma capacidade de agir (*agency*) e resistência políticas que não podem ser descritos exatamente como “vida nua”. Penso que, essas vidas estão saturadas de poder, mesmo que elas estejam excluídas da “*polis*”. Se o campo do poder no qual elas vivem implica certamente uma submissão, essa submissão não é um atributo essencial ou exaustivo. Nós observamos redes de saúde, práticas de mobilização política e formas de resistência em todos esses lugares; e nós devemos, portanto, pensar um modelo de poder que dê conta da diversidade do que existe e do que se produz lá. O Estado não age sempre por um poder “soberano” em sua relação com o povo, visto que a soberania, em certa medida, disseminou-se na governamentalidade. Eu também sugeri que a concepção de federalismo em Hannah Arendt (como aquela que ela propõe para a Palestina) dependesse fundamentalmente de uma distribuição de efeitos soberanos. Inquieto-me com as posições que acentuam o poder central da soberania sobre e à custa do campo das “vidas nuas”. Essas posições são talvez românticas e sedutoras, mas elas não nos ajudam a pensar a formação contemporânea da soberania, nem os modos de investimento e capacidade de agir (*agency*) políticos fora daqueles que estão confinados na *polis* ou excluídos de suas fronteiras. Considero, entretanto, que a noção de vidas “abandonadas” em Agamben é muito útil para reflexão sobre as populações precárias, mesmo que este não seja seu vocabulário.

La Vie des Idées: Atualmente você trabalha muito sobre o caráter determinante dos afetos no campo político, estudando, por exemplo, as imagens de Abu-Graïb. De modo geral, muitos trabalhos atuais mostram a dimensão política dos afetos, tais como a vergonha, a aversão ou o nojo, e renovam, assim, nossa maneira de conceber as relações de poder. Parece, contudo, que as emoções positivas – aquelas que não são “tristes”, como diria Spinoza – constituem o objeto de uma

atenção menos prestigiada. Estes afetos alegres tem uma dimensão política? Que lugar lhes dar na analítica do poder?

Judith Butler: Na realidade, lembrei-me, recentemente, das formas de “expropriação”, até mesmo de “êxtase”, que se produzem nas manifestações de rua; e também me interesse bastante pelas formas de vulnerabilidade que conduzem a uma paixão habitável, e não somente à exploração. Argumentando contra as formas burguesas de casamento, continuo a advogar por um campo da sexualidade desregulamentada.

La Vie des Idées: Na primeira entrevista de *Humain, inhumain* (*Humano, inumano*), você mencionou muito brevemente suas origens judaicas e sua impossibilidade de escrever sobre o tema do judaísmo em relação “à dor e à vergonha que suscita em mim o Estado de Israel”. O conflito israelita-palestino é abordado em seu trabalho atual como uma situação exemplar para compreender as relações de poder e a soberania. Seu novo livro, *Parting Ways: Jewishness and the Critic of Zionism*, é consagrado à questão do judaísmo. O que lhe levou a retomar essa questão? É uma atualidade política?

Judith Butler: Venho de um meio judeu bastante praticante e tentei, com esse livro, voltar-me para minha própria formação, sobre o que me foi ensinado, e fazer o trabalho necessário para avaliar de forma crítica minha educação sionista. Minha crítica ao sionismo está em andamento há algumas décadas, principalmente em minhas conversas reservadas. Mas os debates públicos que seguiram o 11 de setembro, parece, obrigaram-me a tornar pública minha posição sobre esse assunto. Para mim, alguns dos valores judaicos que me foram ensinados – o significado do luto em público e com os outros, a brevidade da vida e, portanto, seu valor, a luta não-violenta – tem trilhado caminhos para se tornarem argumentos mais gerais, até mesmo argumentos contra o sionismo político contemporâneo.

Por essa razão, não creio que se deva renunciar ao seu judaísmo para se opor ao Estado de Israel e não penso que se criticamos esse Estado sejamos, de uma forma ou de outra, antijudeus ou antisemitas (mesmo que seja, por vezes, o caso). Meu objetivo é utilizar minha formação para desenvolver uma série de posições que afirmam a possibilidade de uma vida ética e política dos judeus com os não judeus. É certamente uma posição diaspórica mas, seguindo Edward Said, penso que pode ser um ponto de partida útil para pensar uma política democrática radical na Palestina.

La Vie des Idées: Em seus trabalhos sobre o gênero, a psicanálise era criticada tanto por sua concepção normativa das identidades de gênero quanto por sua representação da mulher. Mas ela foi também uma referência indispensável para pensar a dimensão psíquica dos processos de subjetivação e a vulnerabilidade do sujeito. Você explicou sobre esse primeiro uso da psicanálise, por exemplo, na primeira entrevista de *Humains, inhumain* (“o gênero como performace”). Hoje, em suas reflexões sobre poder, precariedade, encontramos sempre uma referência importante à psicanálise, particularmente através da questão do luto. O que é que hoje, para você, é indispensável na teoria psicanalítica para pensar a política?

Judith Butler: Eu creio que é necessário tentar compreender como e porque os Estados e instituições públicas “repudiam” a interdependência dos seres humanos e porque os numerosos agentes progressistas consideram a dependência propriamente dita, como uma idéia “incontrolável”. É comum o uso da dependência ao serviço de políticas coloniais e paternalistas, enquanto que a interdependência sugere igualdade. Em minha opinião, o sujeito autônomo e sem necessidades, aquele que nunca foi alimentado ou educado por outra pessoa, é uma concepção extremamente problemática do sujeito. Porque este busca se proteger, ele rompe os vínculos sociais e pode sozinho se preservar através da negação e da destruição. É por isso que temos necessidade da psicanálise, mas também, talvez,

de Hegel. Da mesma forma, quando as populações destruídas não são lamentadas, e quando as chamamos de “efeitos colaterais” ou de outros termos desse tipo, há uma negação da violência como perda que deve ser recuperada e combatida.

La Vie des Idées: Ao longo de seus trabalhos, você é cautelosa ao identificar as interdependências que constituem uma vida, [...] o que a torna precária, o fato de ser por definição, em seu ser e em sua identidade, tributária e dependente dos outros. Esse quadro social e intersubjetivo da subjetividade não ocorre alheio a sua relação com Hegel que, mais do que qualquer outro, você destaca em *Sois mon corps (seja meu corpo)*, não pára de recordar que o sujeito está constitutivamente ligado ao todo e ao comum, constituindo seu sentido. Não há nenhum sentido em pensar os fenômenos da subjetividade independentes desse quadro social? Por exemplo, uma relação ética solitária de si para si, independente das interações e interdependências que fazem noutro lugar uma vida.

Judith Butler: Certamente, as relações de si para si existem, mas mesmo quando esse si solitário tende a se passar por objeto de reflexão, ou mesmo tomar conta de si, ele lida com uma série de convenções, de termos e de normas das quais ele próprio não é o autor. São convenções sociais que nos chegam através da linguagem e de um amplo campo de significações sociais, no qual nós todos fomos formados. Quando começamos a refletir sobre nós mesmos, nós não abandonamos essa formação social. Ela está presente nos interstícios de nosso pensamento, e mesmo em nossa concepção do que um “si” deveria ser. Desta forma, embora seja possível estar totalmente isolado em seu pensamento, ou só fisicamente, quando nenhum barulho da rua seja perceptível e que ninguém esteja à vista, o traço vivo do mundo social continua a mediar as relações mais íntimas que mantemos com nós mesmos.

Entrevista traduzida do inglês para o francês por Barbara Turquier.

Vá mais longe – Principais obras de J. Butler, traduzidas em francês:

La Vie psychique du pouvoir. L'Assujettissement en théories, Paris, Leo Scheer, 2002.

Antigone. La Parenté entre vie et mort, Paris, EPEL, 2003.

Le Pouvoir des mots. Politique du performatif, Paris, Amsterdam, 2004.

Vie précaire. Les Pouvoirs du deuil et de la violence après le 11 septembre 2001, Paris, Amsterdam, 2005.

Humain, Inhumain. Le Travail critique des normes. Entretiens, Paris, Amsterdam, Paris, 2005.

Trouble dans le genre. Pour un féminisme de la subversion, Paris, La Découverte, 2005.

Défaire le genre, Paris, Amsterdam, 2006.

Le Récit de soi, Paris, Puf, 2007.

L'État global, avec Gayatri Chakravorty Spivak, Paris, Payot et Rivages, 2007.

Ces corps qui comptent ; de la matérialité et des limites discursives du « sexe », Paris, Amsterdam, 2009.

Sois mon corps, avec Catherine Malabou, Paris, Bayard, 2010.

Ce qui fait une vie, Paris, Zone/La Découverte, 2010.

Sujets du désir, réflexions hégéliennes en France au xxe siècle, Paris, PUF, 2011.

Para citar este artigo: Claire Pagès & Mathieu Trachman, «Une analytique du pouvoir. Entretien avec Judith Butler». *La Vie des idées*, 4 décembre 2012. ISSN : 2105-3030. <http://www.laviedesidees.fr/Une-analytique-du-pouvoir.html>