

# AUTOCONSCIÊNCIA E PREDICAÇÃO DE SI

*Roberto Horácio de Sá Pereira*

UFRJ/CNPq

## INTRODUÇÃO

O problema central da autoconsciência emerge a partir do que do que Henrich denominou “teoria da reflexão” <Reflexionstheorie> (Henrich, 1967). Sua característica essencial consiste na concepção do fenômeno da autoconsciência como o resultado de um ato do pensamento que se volta sobre o próprio sujeito de tal ato <Sichzurückwenden> tomando-o como seu objeto. A mesma suposição fundamental sobrevive na filosofia contemporânea da mente sob a forma da chamada teoria “do pensamento de ordem superior” (HOT). Um sujeito só tornaria consciente de si mesmo quando, por meio de um pensamento de ordem superior (HOT), identificasse ou apreendesse a si mesmo como o autor ou o sujeito pensante de algum pensamento de ordem inferior (conferir Rosenthal 2004). Deste modo, tomar consciência de si seria apreender a si mesmo como um objeto perfilado diante da mente.

O primeiro a tomar ciência dos problemas gerados por tal teoria foi Fichte. Se a autoconsciência for entendida como o resultado de um ato de reflexão, a referência consciente a si mesmo não seria bem sucedida a menos que o indivíduo

que reflete *já soubesse* de algum modo que ele próprio é quem está realizando o ato de reflexão. Assim o teórico da reflexão se veria às voltas com o seguinte dilema. Ou bem seria necessário um ato de reflexão de segunda ordem (no sentido de que eu sou o indivíduo realizando o ato de reflexão original) e esse ato exigiria, por sua vez, um outro ato de ordem superior e, assim, indefinidamente. Ou bem a consciência de si seria dada no próprio ato de reflexão. Mas como a consciência de si deve resultar da realização do ato de reflexão, o sujeito da reflexão não teria como saber que ele próprio estaria realizando tal ato antes ou durante a sua própria realização.

Fichte imaginava poder contornar tal problema, concebendo a autoconsciência como uma forma imediata de familiaridade consigo mesmo <Kenntnis seiner selbst>. Tomaríamos consciência de nós mesmos na medida em que, ao refletirmos nos “auto-colocamos” <Sichselbstsetzen> como sujeito e objeto do nosso próprio pensamento. Ao que tudo indica, no entanto, a solução proposta por Fichte se vê às voltas com o mesmo problema que pretendia resolver. Segundo Pothast, o mais influente discípulo de Henrich nos anos setenta, não é absolutamente compreensível como o eu poderia se auto-colocar a partir de um ato de reflexão levada a cabo por si mesmo. Ademais, acrescenta Henrich, a circularidade viciosa não é eliminada ao qualificarmos simplesmente a auto-referência como uma relação imediata. Só posso me colocar reflexivamente como um objeto para meu próprio pensamento se já sei que sou o indivíduo realizando tal colocação. A solução encontrada pela escola de Heidelberg consiste na reiteração da posição defendida por toda a tradição fenomenológica que reside na interpretação da familiaridade imediata consigo mesmo como uma forma pré-reflexiva de consciência de si na qual não haveria mais a distinção entre sujeito e objeto, entre a experiência refletida e a experiência da reflexão.

A abordagem analítico-lingüística de Tugendhat surge como o resultado da sua avaliação crítica da posição tradicional. Segundo tal avaliação, o problema fundamental levantado por Fichte e Henrich só tem lugar uma vez que a autoconsciência é pensada nos termos do modelo tradicional sujeito-objeto segundo o qual o sujeito põe ou apreende a si mesmo como um objeto perfilado <Gegenstand> *diante*

da mente <Vorstellung>. Colocando de lado a teoria da reflexão, a autoconsciência passa a ser compreendida como o conhecimento imediato de si que teria a forma de uma simples proposição em primeira pessoa <Ich-Sätze>, governada por dois princípios fundamentais: o princípio da simetria veritativa (entre predicções na primeira e terceira pessoas fazendo referência a um mesmo estado de coisas) e o princípio da assimetria epistêmica (entre o conhecimento de si na primeira pessoa e o mesmo conhecimento na perspectiva de terceiros).

Esse trabalho foi concebido como uma avaliação crítica da posição de Tugendhat. Nele pretendo sustentar, em linhas gerais, a posição lingüístico-analítica de Tugendhat segundo a qual a autoconsciência se estrutura fundamentalmente sob a forma de uma simples predicção de si. Como Tugendhat, rejeito, portanto, a teoria tradicional da reflexão segundo a qual a auto-referência reflexiva surgiria como resultado de um pensamento de ordem superior (HOT) que identificaria o sujeito como o autor de pensamentos ordinários. Mas mesmo rejeitando tal modelo, buscarei mostrar que a crítica de Tugendhat não é convincente por diferentes razões.

O trabalho está concebido da seguinte forma. Na primeira seção, apresento a posição analítica de Tugendhat como resultado de sua crítica à posição tradicional, representada na filosofia contemporânea pela escola de Heidelberg. Na seção seguinte, empreendo uma avaliação da crítica de Tugendhat à posição tradicional, buscando assinalar as diferentes razões pelas quais ela não me parece satisfatória. Ao final, apresento a minha própria explicação para a auto-referência cognitiva e o auto-conhecimento imediato a partir da semântica informacional.

#### A CONCEPÇÃO ANALÍTICO-LINGÜÍSTICA DE TUGENDHAT

Segundo o diagnóstico de Tugendhat, o problema tradicional de Fichte só surge porque a autoconsciência é entendida nos termos tradicionais do modelo sujeito-objeto. Segundo esse modelo epistemológico tradicional (ainda vigente na filosofia da mente contemporânea), objetos estariam perfilados diante da mente que os apreenderia ou por uma percepção interna ou por um pensamento de ordem superior. Tomar consciência de si seria então *por* <stellen>, ou seja, apreender a si

mesmo como um objeto perfilado *diante* <vor> da mente <Vorstellung>. Tugendhat resume suas críticas à posição tradicional nos seguintes termos:

A dificuldade com a teoria da reflexão identificada por Henrich da qual Fichte partia se nutre da suposição de que se trata de algo cuja essência consiste na identidade entre conhecer e ser conhecido. Para alguém que não reconheça que o fenômeno da autoconsciência tenha ou pressuponha tal estrutura, a dificuldade não existe. A dificuldade, que é, com efeito, insolúvel, é apenas o reflexo do absurdo de tal abordagem (Tugendhat, 1979, 64).

Mas Tugendhat não se limita a rejeitar a teoria tradicional da reflexão. Ele recusa também a solução fenomenológica para o problema da circularidade que postula uma forma pré-reflexiva e pré-lingüística de consciência de si mesmo <Kenntnis seiner selbst> e que estaria presente tacitamente em toda experiência. Segundo Tugendhat, toda consciência intencional (toda consciência *de* algo) exibiria uma forma proposicional. Segundo sua apreciação, o verbo germânico “estar familiarizado com” <kennen> é sempre empregado de forma proposicional. À primeira vista, tal verbo transitivo direto exige apenas um objeto como completo e não uma proposição, como a maioria dos verbos de atitudes proposicionais. Um exame mais detido revela, contudo, que as locuções “estou familiarizado com algo” e “estou familiarizado comigo mesmo” <Ich kenne mich> constituem elipses. Ao afirmar “estou familiarizado comigo” suscito imediatamente uma pergunta sobre o que estou afirmando conhecer de mim mesmo. Portanto, quem afirma “estou familiarizado comigo” estaria dizendo qualquer coisa como “sei que sou assim e assim” <Ich kenne mich als so und so> (conferir, Tugendhat 1979, 57-8).

A consciência pré-reflexiva e pré-lingüística se reduziria assim a uma mera sensação destituída de conteúdo ou referência. Animais e crianças pequenas não teriam nem consciência *de* objetos nem consciência de si, mas apenas sensações sem conteúdo. Com a linguagem proposicional, surgiriam ao mesmo tempo a consciência intencional de nós mesmos e a consciência intencional dos objetos que nos cercam. Não teríamos como tomar consciência de nós mesmos como um

dentre outros objetos que se encontra em um determinado estado a menos que tomássemos consciência dos objetos que nos cercam. E, inversamente, não teríamos como tomar consciência dos objetos que nos cercam ao pensá-los como um dentre outros objetos senão por referência à consciência que possuímos de nós mesmos. No seu trabalho mais recente, Tugendhat resume essa tese nos seguintes termos:

Contudo, na minha opinião, a “mera consciência” não é uma consciência de objetos, senão o fenômeno pré-lingüístico dos estados mentais que não possuem mais do que a propriedade da consciência que ainda não é consciência *de*. Com a linguagem predicativa, surgem ao mesmo tempo a consciência de *outros* objetos e a consciência de *si mesmo* como um objeto dentre outros (...). O indivíduo que fala uma linguagem proposicional não poderia ter consciência de si mesmo a menos que tivesse consciência do todo – de um mundo objetivo -, e não poderia ter consciência de um mundo objetivo se não pudesse fazer referência a si mesmo (Tugendhat, 2004, 32).

Segundo a interpretação crítica de Tugendhat, o problema da circularidade só se coloca quando a autoconsciência é erroneamente entendida como a suposta consciência imediata da identidade entre o sujeito da representação e o sujeito representado, resultante do ato de reflexão. Se, por um lado, a consciência imediata da identidade de si deveria ser estabelecida no próprio ato de reflexão, por outro lado, o próprio ato de reflexão não poderia ser realizado a menos que o sujeito da reflexão já estivesse cômico da sua própria identidade. Assumindo a consciência prévia da identidade do sujeito, a teoria da reflexão acaba por pressupor o que deveria antes ser explicado pelo próprio ato da reflexão (Tugendhat, 1979, 62). Mas na medida em que o fenômeno da autoconsciência imediata não se estrutura sob a forma da consciência de uma relação de identidade de si para consigo mesmo, mas, mas antes como uma predicação sobre si mesmo, a conclusão que se impõe é que o círculo vicioso de Fichte não passa de um problema sem sentido <Sinnwidrigkeit>.

Assim, se a autoconsciência em geral toma a forma de uma de uma predicação de si <Ich-Sätze>, a chamada autoconsciência imediata exibe a forma par-

cular de uma proposição psicológica em primeira pessoa na qual o predicado auto-atribuído é mental. Tendo descartado a teoria tradicional da reflexão, a abordagem analítico-lingüística de Tugendhat deve ainda responder a duas questões fundamentais. Ela deve esclarecer, em primeiro lugar, o emprego do pronome de primeira pessoa como um termo singular peculiar no contexto do pensamento em primeira pessoa. Ademais, ela deve tornar compreensível o próprio estatuto imediato da autoconsciência.

Como último ponto de referência, o pronome de primeira pessoa não identifica minha própria pessoa como um dentre outros indivíduos em um domínio. Contrariamente a Wittgenstein e Anscombe, Tugendhat sustenta que o pronome de primeira pessoa em uma predicação mental em primeira pessoa se refere à minha pessoa como um *indivíduo passível de ser identificado* a partir da perspectiva de uma terceira pessoa. Assim, Tugendhat estabelece o que denomina princípio da simetria veritativa entre predicacões psicológicas em primeira e terceira pessoas do singular:

A frase <Satz> “eu  $\varphi$ ”, quando proferida por mim, é *necessariamente* verdadeira exatamente quando a frase “ele  $\varphi$ ” for verdadeira, quando proferida por uma terceira pessoa a qual por meio do emprego do pronome “ele” se refere à minha pessoa. (1979, 88, grifo do autor).

Quando alguém se refere a si mesmo por meio do seu nome próprio e uma terceira pessoa se refere a um mesmo indivíduo por meio de uma descrição, duas diferentes proposições estão sendo expressas à luz da teoria clássica da proposição (Frege). Isso porque quem não saiba da coreferencialidade dos termos singulares envolvidos pode, de forma consistente, afirmar uma das predicacões e negar a outra ou vice-versa. Suponhamos que ao investigar o assassinato de Laio na condição de rei de Tebas, Édipo seja levado a conjecturar o seguinte:

(1) O assassino de Laio sente culpa pelo seu feito.

Quando uma terceira pessoa pensasse:

(2) Édipo sente culpa pelo seu feito,

estaria se referindo à mesma pessoa. Mas na medida em que Édipo não soubesse que nome “Édipo” e a descrição definida “o assassino de Laio” se referem a uma mesma pessoa (ele próprio), Édipo tanto poderia de forma consistente tanto afirmar (1) e negar (2), quanto afirmar (2) e negar (1).

Segundo a teoria clássica da proposição, (1) e (2) exprimem proposições ou estados de coisas diferentes. Na avaliação de Tugendhat, a coisa muda de figura quando os termos singulares envolvidos são as expressões dêiticas <indexicals> “eu” e “ele”. Quando alguém se refere a si mesmo por meio do pronome de primeira pessoa e um terceiro (ou a própria pessoa) se refere ao mesmo indivíduo por meio do pronome de terceira pessoa, uma mesma proposição estaria sendo expressa. Assim, Édipo não poderia afirmar (3) e negar de forma consistente (4):

(3) Ele (fazendo referência à sua própria pessoa) se sente culpado.

(4) Sinto-me culpado.

Entretanto, embora uma mesma proposição esteja sendo expressa a partir da perspectiva da primeira e terceira pessoas, apenas aquele que entretém o pensamento (4) possuiria um conhecimento *imediato* de si. São observações e inferências que levam alguém a pensar em (3). Assim, a simetria semântica é seguida por uma assimetria epistemológica entre as orações psicológicas em primeira e terceira pessoas. Em outras palavras, a mesma proposição é conhecida de forma distinta a partir da perspectiva da primeira e da terceira pessoas. Para conhecer a verdade de (3), tenho que observar a conduta da pessoa a quem atribuo culpa ou fazer inferências, enquanto que para que conhecer a verdade de (4) nenhuma observação ou percepção interna do estado mental em questão seriam necessários.

Mas se o suposto conhecimento imediato de si enquanto sujeito de estados mentais não pode ser compreendido como a percepção interna de tal estado mental, Tugendhat nos deve uma explicação sobre como entendê-lo positivamente. Seguindo Wittgenstein, Tugendhat afirma que proposições psicológicas em primeira pessoa são imunes a um erro de referência quando empregadas em conformidade com a sua própria regra de emprego. Assim, se compreendo a regra do emprego

do pronome de primeira pessoa, não posso deixar de reconhecer que ao empregar tal pronome estou fazendo referência à minha própria pessoa. Da mesma maneira, se compreendo a regra de uso do predicado “sentir culpa”, não posso deixar de reconhecer tampouco que o que sinto é culpa (quando estou sentindo alguma coisa).

Tendo reconhecido que o emprego do pronome “eu” é referencial, Tugendhat também se recusa a assimilar as proposições psicológicas em primeira pessoa a meras expressões de estados mentais, como propunha Wittgenstein. Embora as predicções psicológicas da forma (4) não possam ser entendidas como asserções, como as predicções correspondentes na terceira pessoa (3), elas ainda assim exprimem uma forma imediata de saber <Wissen seiner selbst> não baseado, contudo, em um ato de cognição (Conferir, 1979, 133). Assim, Tugendhat conclui que ao empregar uma proposição psicológica em primeira pessoa em conformidade com a sua própria regra, o sujeito não pode deixar de reconhecer sua verdade de forma imediata.

**AUTO-CONHECIMENTO IMEDIATO E AUTO-REFERÊNCIA IMEDIATA. A PRIMEIRA QUESTÃO** a ser levantada diz respeito à crítica da idéia da tradicional do conhecimento imediato de si como uma forma de familiaridade consigo mesmo <Kenntnis seiner selbst>. Como observamos, Tugendhat rejeita a postulação tradicional de uma forma pré-reflexiva e pré-lingüística de consciência de si alegando, primeiro, que toda consciência intencional (toda consciência *de* algo) exibiria uma forma proposicional e, segundo, que a consciência intencional dos objetos que nos cercam pressuporia uma referência a nós mesmos como sujeitos das nossas experiências. Para Tugendhat, uma consciência pré-reflexiva e pré-lingüística se reduziria a uma série de sensações destituídas de conteúdo.

Essa argumentação não é convincente por inúmeras razões. Em primeiro lugar, a alegação de que locutor da expressão “estou familiarizado comigo mesmo” suscita a questão “o que você sabe sobre você mesmo?” (conferir, Tugendhat, 1979, 46) incorre em uma petição de princípio. Ao tomarmos consciência *de* algo não estamos necessariamente fazendo uma asserção ou entretendo um pensamento sobre um determinado objeto. Ao afirmarmos, por exemplo,



(5) O cão está familiarizado com o osso.

Não estamos atribuindo ao animal uma atitude proposicional que representaria a entidade em questão *como* um osso. Estamos antes atribuindo ao cão uma atitude intencional frente a uma determinada entidade independentemente de como essa entidade possa ser caracterizada por meio de conceitos. Afinal, o que mais nos permitiria explicar o complexo comportamento intencional dos animais de solucionar problemas? Não temos como tornar compreensível o comportamento de um cão que se põe a cavar onde um osso encontra-se escondido, atribuindo-lhe apenas uma sensação sem conteúdo. A postulação de conteúdos não-conceituais se sustenta fundamentalmente no princípio da melhor explicação disponível da conduta observada (conferir Bermúdez 1997). O comportamento do cão só se torna compreensível sob a suposição de que ele tem consciência intencional *do* osso sendo capaz, ademais, de se recordar de tal objeto quando não o percebe no seu campo visual.

A experiência pré-lingüística e pré-reflexiva se caracteriza, portanto, por uma série de atitudes intencionais *de re* (diante da própria coisa) que seriam independentes de atitudes proposicionais *de dicto* que representariam as mesmas entidades por meio de conceitos. Perceber e estar familiarizado *com* um determinado objeto não significa ter consciência de tal objeto como próprio à extensão de um conceito. Nem significa tampouco possuir a consciência proposicional (i) de que se está realizando uma experiência determinada e (ii) de que tal experiência *lhe* pertence. Para tanto, além da capacidade de conceituar o que percebe, o indivíduo também deveria ser capaz de conceituar a própria experiência que realiza como um experiência específica e, ademais, de caracterizá-la como *sua*. Mas o ponto fundamental é o seguinte. Mesmo incapaz de conceituar o que percebo, de conceituar a experiência que realizo e de conceituar a mim mesmo como seu autor, não deixo de perceber algo determinado e de tornar-me familiarizado tanto com a experiência que realizo quanto comigo mesmo como seu autor. Em oposição, entretanto, à tradição fenomenológica, que toma as noções de “consciência” e familiaridade consigo mesmo <Kenntnis seiner selbst> como um dado imediato e irreduzível, buscaremos uma explicação para essa forma não-conceitual de consciência de si a partir da teoria da informação.

A identificação da autoconsciência imediata com uma forma imediata de auto-*conhecimento* em sentido próprio levanta uma série de outros problemas. Segundo Tugendhat, se conheço o significado ou a regra do emprego de uma determinada predicação psicológica (4) e me encontro em um estado mental qualquer, então tenho que *saber* que eu me encontro no estado mental em questão. Nestes termos, mesmo excluindo a idéia tradicional de uma percepção interna, Tugendhat permanece subscrevendo algumas teses tradicionais acerca da autoconsciência. A mais conhecida é certamente a tese de que proposições psicológicas em primeira pessoa são sempre verdadeiras. Se me encontro em algum estado mental e conheço a regra do emprego de da oração (4), então posso ter certeza de que a minha crença (4) tem que ser *verdadeira*. Nas palavras de Tugendhat:

Já observamos que a frase <Satz> “eu  $\varphi$ ” também é verdadeira quando empregada meramente em conformidade com sua regra. Disso se segue, contudo, que eu também posso ter certeza <Gewissheit> de que a frase “eu  $\varphi$ ” é verdadeira quando a emprego meramente em conformidade com a regra (conferir, Tugendhat, 1979, 133).

Como tal crença apresenta uma estrutura predicativa, a certeza em questão pode ser relativa tanto ao componente predicativo quanto ao elemento da identificação. Assim, sabendo o que o predicado “sentir culpa” significa, não haveria possibilidade de que eu pudesse representar erroneamente o estado mental no qual me encontro como sendo um estado de culpa. E conhecendo a regra de emprego do pronome de primeira pessoa, também não haveria a possibilidade de que eu pudesse fazer uma referência errônea à minha pessoa como aquela que está sentindo culpa.

Na tradição empirista, seria uma percepção interna de mim mesmo e do meu próprio estado mental que explicaria porque minha crença predicativa não pode ser corrigida. Para Tugendhat, contudo, a certeza repousa no conhecimento da regra de emprego do predicado e na regra de emprego do pronome de primeira pessoa. Uma vez que o conhecimento é visto tradicionalmente como uma crença verdadeira e justificada, ao pensar em (4) e conhecendo as respectivas regras

envolvidas, segue-se então que *sei* imediatamente (sem observação ou percepção interna) que estou sentindo culpa. A autoconsciência seria um conhecimento imediato de si.

Poucos filósofos hoje em dia parecem dispostos a defender a infalibilidade da caracterização dos nossos estados por meio de predicados mentais. Com base em contra-exemplos, é relativamente fácil mostrarmos como a caracterização dos nossos próprios estados por meio de predicados mentais pode incorrer em erro, mesmo quando os predicados mentais envolvidos são relativamente simples, como o predicado “sentir dor”. Um dos contra-exemplos mais ilustrativos vem dos rituais de iniciação. Tendo seus olhos vendados, é dito a um indivíduo que uma navalha está prestes a cortar um determinado ponto do seu pescoço. Entretanto, no lugar da navalha, um cubo de gelo é colocado no ponto indicado. Curiosamente, nesse instante, a grande maioria dos indivíduos testados reporta sentir dor no local indicado, quando na verdade experimentam a sensação de frio. Assim, mesmo sabendo o que significa “dor” (ou seja, conhecendo a regra de emprego do predicado psicológico “dor”), e sentido efetivamente alguma coisa (por exemplo, sensação do gelo), podemos ainda acreditar erroneamente que dor é o que estamos sentindo, mesmo que auto-atribuição do predicado “dor” independa de uma suposta percepção interna, de uma observação da própria conduta ou de uma inferência.

Um modo de contornar essa dificuldade consiste em abrandar a pretensão à infalibilidade. Ao invés de afirmar que sempre que creio que sinto dor (e sei o que significa afirmar “tenho dor”), é verdade que dor é o que sinto, se poderia afirmar que sempre que creio estar com dor é verdade pelo menos que me encontro em *algum* estado mental. Assim, se ao acreditar que estou com dor posso não saber que dor é o que eu sinto, sei pelo menos (e de forma imediata) que me encontro em algum estado mental. Talvez essa infalibilidade predicativa nuançada possa contornar a dificuldade levantada. Entretanto, é importante salientar que a auto-atribuição de tal predicado só se torna infalível porque não veicula mais nenhuma informação substantiva sobre a natureza do estado mental no qual o sujeito se encontra. Afinal, o que significa se afirmar que sei que me encontro em um estado mental se não sou capaz de saber precisamente de qual estado se trata?

Tugendhat não está sozinho, no entanto, ao pensar que um erro de identificação estaria excluído da autoconsciência entendida como uma predicação sobre si mesmo. Wittgenstein foi certamente o primeiro a tomar ciência dessa peculiaridade semântica ao distinguir dois tipos de orações em primeira pessoa. O primeiro seria aquele no qual o pronome de primeira pessoa é usado *como sujeito*. Exemplos seriam orações da forma:

(6) Tenho dor

O segundo tipo é caracterizado pelo emprego do mesmo pronome *como objeto*. Tal emprego é ilustrado por Wittgenstein em orações como:

(7) Tenho o braço quebrado.

O que distingue o emprego do pronome de primeira pessoa como sujeito do emprego do mesmo pronome como objeto é a impossibilidade de um erro peculiar. Enquanto o emprego de tal pronome como objeto envolve o *reconhecimento* de uma pessoa dentre outras e está, por essa razão, sujeito a um erro de identificação, o emprego do pronome de primeira pessoa como sujeito não estaria sujeito ao mesmo tipo de erro uma vez que não faria referência a nada (Wittgenstein 1958).

Shoemaker acolhe a distinção proposta por Wittgenstein, mas sustenta contra este que em seu emprego como sujeito o pronome de primeira pessoa é referencial, embora de uma forma diferente do seu emprego como objeto (Shoemaker 1968). Se podemos inferir que alguém tem dor da afirmação (6) como também que duas pessoas têm dor da afirmação (6) e da suposição de que Paulo tem dor, o pronome da primeira pessoa possui inegavelmente a função de termo singular em (6).

Shoemaker busca tornar compreensível a distinção proposta por Wittgenstein, analisando as orações em primeira pessoa nos seus dois componentes básicos: a identificação, realizada pelo termo singular “eu”, e a predicação, realizada por um termo geral. No primeiro tipo de conteúdo (o eu como objeto), estamos sujeitos a dois tipos de erro: um relativo ao componente predicativo e o outro relativo ao componente da identificação. Acreditando, por exemplo, em (7) posso corretamente identificar um objeto (meu braço) que não esteja quebrado efetivamente. Mas tendo corretamente acreditado que alguém tem o braço quebrado,

posso supor erroneamente que é o *meu* braço que está quebrado quando se trata de um braço alheio.

Já no segundo tipo de conteúdo (o eu como sujeito), estamos sujeito ao erro apenas relativamente ao componente predicativo. Ao pensar em (6), posso me equivocar quanto ao fato de ser dor o que estou sentido, mas jamais quanto ao fato de ser eu o indivíduo que está sentido dor. A impossibilidade de tal erro foi batizada por Shoemaker como a “imunidade ao erro por identificação relativo ao emprego do pronome da primeira pessoa do singular”.

Shoemaker concebe duas formas de imunidade ao erro por identificação. A primeira seria uma imunidade circunstancial manifesta em orações da forma:

(8) Estou diante de uma mesa.

Em situações usuais ao afirmar (8) posso me equivocar quanto ao fato de ser uma mesa o objeto diante do qual eu me encontro, mas não quanto ao fato de ser eu que me encontro diante de tal objeto. Há circunstâncias, contudo, nas quais ao afirmar (8) eu poderia erroneamente tomar uma terceira pessoa por mim mesmo (quando, por exemplo, eu fizesse tal afirmação (8) observando um espelho). (Conferir, Shoemaker, 1968, 82). Tal imunidade circunstancial ao erro por identificação seria derivada, entretanto, de uma forma absoluta de imunidade ao erro por identificação que poderia ser representada nos termos da seguinte oração:

(9) Vejo uma mesa no centro do meu campo visual.

Se há circunstâncias nas quais a afirmação (8) pode estar baseada em erro de identificação, ao afirmar (9) jamais poderia me equivocar quanto ao fato de ser eu quem percebe a mesa no centro do seu campo visual. Uma vez que a auto-atribuição expressa por (8) deve ser vista como uma consequência indutiva da auto-atribuição expressa por (9), segundo Shoemaker, o que confere imunidade circunstancial a (8) é a imunidade absoluta expressa por (9). O contraste entre (8) e (9) nos permite compreender a origem da imunidade ao erro por identificação: enquanto em (8) o sujeito se auto-atribui um predicado não-mental (estar diante de uma mesa), em (9) o sujeito se auto-atribui um predicado mental (perceber uma mesa), nas palavras de Shoemaker, um predicado que não poderia ser ins-

tanciado por mim sem que eu soubesse que eu estou a instanciando (Conferir, 1968, 90).

A posição comum a Shoemaker e Tugendhat é objeto de duas objeções fundamentais. Em primeiro lugar, como Evans corretamente assinala, a imunidade ao erro por identificação não parece estar restrita à classe das proposições psicológicas do em primeira pessoa na quais um predicado mental é auto-atribuído (conferir Evans, 1982, capítulo 7). Suponhamos a seguinte afirmação:

(10) Minhas pernas estão cruzadas.

Sob a suposição de que sei que as pernas de alguém estão cruzadas, não faz sentido indagar se sei que são as *minhas* pernas que estão cruzadas e não as pernas de uma outra pessoa. Mas se é assim, a ausência de identificação na predicação de si não pode ser explicada pela auto-atribuição de um predicado mental, na forma como supunham Shoemaker e Tugendhat. A tese central de Evans é a de que a imunidade ao erro por identificação não seria uma propriedade de proposições *simpliciter*, mas antes de juízos ou crenças sobre proposições a partir dos seus modos de justificação. Assim o juízo expresso por (6) será imune ao erro por identificação quando sua justificação estiver baseada nas sensações de dor do próprio sujeito. Da mesma forma, o juízo expresso por pela asserção (10) também estará imune ao erro por identificação quando sua justificação estiver baseada em informações proprioceptivas sobre a posição dos membros do próprio corpo. Em ambos os casos, a informação de que uma determinada propriedade está sendo instanciada é *normalmente* acompanhada pela informação adicional de que sou eu quem a estou instanciando.

Ademais, reconhecer com Shoemaker e Tugendhat a ausência de identificação em determinadas predicções de si não significa reconhecer a suposta imunidade absoluta a um erro de referência relativo ao emprego do pronome da primeira pessoa, conferida *a priori* a tais predicções pelo emprego de um predicado mental e pela regra de uso de tal pronome. A esse respeito, Rosenthal nos convida a imaginar a seguinte situação anormal. Suponhamos que eu tenha a propensão patológica de estabelecer empatia com o sofrimento alheio e me encontre na situa-

ção na qual tenha a informação de que alguém na minha vizinhança está com dor. Confuso, passo então a acreditar em (6) quando a bem da verdade não sou eu, mas uma terceira pessoa que sente dor.

O único erro que não posso cometer é o de supor que o indivíduo que *tomo* como estando com dor não é aquele que está entretendo esse mesmo pensamento sobre si. Isso é o que Rosenthal denomina imunidade fraca <thin immunity> ao erro por identificação (conferir, Rosenthal 2004). Embora possa me equivocar ao pensar que sou o indivíduo com dor (quando alguém efetivamente sente dor), não posso me equivocar ao supor que o indivíduo que entretém tal predicação de si é o mesmo que em tal predicação é representado como estando com dor.

A conclusão que se impõe é que não há nenhuma regra semântica ou razão *a priori* que torne uma predicação de si absolutamente imune ao erro de referência e ao erro predicativo. Quanto mais informação uma predicação de si veicular sobre a condição na qual o sujeito se encontra, mais sujeita estará a um erro predicativo. E quanto mais informação uma predicação de si veicular sobre o sujeito que se encontra em um determinado estado, mais sujeita estará a um erro de referência. Uma predicação de si só será absolutamente imune ao erro predicativo quando, no caso limite, não veicular nenhuma informação substantiva sobre o estado ou condição que atribui a um sujeito: “creio que me encontro em algum estado indeterminado”. E só será absolutamente imune ao erro de referência quando, no caso limite, não veicular nenhuma informação substantiva sobre o indivíduo ao qual é atribuído tal estado determinado: “creio que sou o indivíduo que é representado por mim mesmo como se encontrando em um determinado estado”.

Duas conseqüências parecem se impor. Primeiro, uma vez que predicções de si podem conter tanto erros predicativos quanto erros de identificação, a auto-referência cognitiva não pode ser identificada a uma forma de conhecimento de si em sentido próprio (crença verdadeira e justificada). Mesmo conhecendo o significado do predicado “dor”, nem sempre que creio ter dor, *sei* que é dor o que eu sinto. E mesmo conhecendo a regra de emprego do pronome da primeira pessoa, nem sempre que creio que sinto dor, *sei* que sou eu quem efetivamente sente dor.

E como a ausência de identificação se explica pelas informações que o sujeito tem de si mesmo, o que Tugendhat denomina autoconsciência epistêmica imediata é independente do emprego de predicados mentais. Sei que são *minhas* as pernas que estão cruzadas ou sei que sou eu quem está sentado escrevendo ao computador da mesma forma imediata como sei que sou eu quem sente dor ou que sei que sou eu quem percebe uma mesa no centro do seu campo visual.

#### AVALIANDO A CRÍTICA DE TUGENDHAT

Como observamos, para Tugendhat, a autoconsciência só constitui um problema quando abordada segundo o modelo sujeito-objeto tradicional. Uma redescritção do fenômeno em termos predicativos seria capaz de contornar o problema da circularidade. Ter consciência de si significa então saber que um determinado predicado mental  $\varphi$  se aplica à pessoa referida (sem ser identificada) pelo pronome de primeira pessoa em um pensamento sobre si. O conteúdo dessa predicação de si é determinado fundamentalmente pelo chamado princípio da simetria veritativa. Segundo esse princípio, a oração psicológica em primeira pessoa “eu  $\varphi$ ” só será verdadeira, quando a oração psicológica em terceira pessoa “ele  $\varphi$ ” também o for quando enunciada para fazer referência a mim mesmo.

Mas mesmo que o pronome pessoal “ele” se refira ao mesmo indivíduo referido pelo pronome de primeira pessoa, há inúmeras situações nas quais alguém poderia afirmar “eu  $\varphi$ ” e negar de forma consistente “ele  $\varphi$ ” ou, inversamente, afirmar “ele  $\varphi$ ” e negar “eu  $\varphi$ ”. No exemplo anterior, Édipo poderia de forma consistente afirmar (3) e negar (4) ou, inversamente, afirmar (4) e negar (3). Suponhamos que o pronome pessoal “ele” se refira a Édipo anaforicamente, por meio da descrição definida “o assassino de Laio” ou por meio do nome próprio “Édipo”:

(11) O assassino de Laio cometeu um crime. *Ele* se sente culpado.

(12) Édipo cometeu um crime. *Ele* se sente culpado.

Não tendo consciência de que ele próprio é o assassino de Laio, Édipo poderia de forma consistente afirmar (4) e negar (11) ou, inversamente, poderia afirmar (11) e negar (4). De forma similar, sofrendo de amnésia, ou desconhecendo que ele



próprio é conhecido como “Édipo”, Édipo poderia de forma consistente afirmar (12) e negar (4) ou, inversamente, afirmar (4) e negar (12).

Mas quando Tugendhat concebe a simetria veritativa entre (3) e (4), ele tem em mente o emprego do pronome da terceira pessoa ou como um substituto <proxy> de uma descrição demonstrativa ostensiva: “essa pessoa” ou simplesmente como um pronome demonstrativo: “esse aí” (Conferir, Tugendhat, 1979, 87):

No seu famoso trabalho da década de sessenta, Castañeda nos apresenta uma série de argumentos contrários à suposta equivalência entre (4) e (3). (Conferir, Castañeda, 1966, páginas 42-45). A experiência realizada por Mach em Viena (retratada por Perry em diferentes trabalhos sobre o tema) nos permite ilustrar mais facilmente o ponto principal. Após um dia de trabalho estafante, Mach (um professor famoso) toma um ônibus de volta para casa. Mas ao adentrar no veículo, observa um professor maltrapilho vindo na direção contrária à sua. Nesse momento Mach pensa consigo mesmo: “Olha como *ele* (essa pessoa) está mal vestido”. Contudo, à medida que se aproxima do seu assento, Mach se dá conta de que ele estava observando sua própria imagem refletida no espelho e que ele, ou seja, *essa pessoa* maltrapilha, a quem ele se referia de forma ostensiva instantes atrás, era ele próprio.

Observando sua conduta no espelho, Édipo poderia de forma consistente afirmar (3) e negar (4) ou, inversamente, afirmar (4) e negar (3), uma vez que ele não teria ciência de que ele próprio é *essa pessoa* a quem está se referindo por meio do pronome da terceira pessoa “ele”. Assim, à luz da teoria clássica da proposição, as orações (3) e (4) exprimem proposições ou conteúdos distintos na medida em que Édipo pode afirmar (4) e negar (4) ou, inversamente, afirmar (3) e negar (4) de forma consistente (tal como no exemplo clássico fregeano as frases “Vênus é a estrela da manhã” e “Vênus é a estrela da tarde” exprimem diferentes proposições <Gedanken>). Segundo o exemplo anterior, mesmo conhecendo o significado dos termos singulares e gerais envolvidos, Mach poderia acreditar de forma consistente que *ele* (ou seja, *esse homem* que ele observa caminhado na direção contrária à sua) está maltrapilho, sem ter que acreditar, contudo, que ele próprio está maltrapilho e vice-versa.

Se quisermos reconhecer a simetria veritativa entre (3) e (4) proposta por Tugendhat, reconhecendo que (3) e (4) enunciam um mesmo conteúdo (ou seja, possuem as mesmas condições de verdade), devemos abandonar a teoria clássica da proposição em favor da teoria referencialista proposta por Kaplan (Kaplan, 1989b). Ao enunciar (3) e (4), Édipo estaria exprimindo, segundo expressão cunhada por Kaplan, uma mesma *proposição singular* que tem o próprio Édipo como seu constituinte e não uma condição da sua identificação (um sentido fregeano <Sinn>) expressa por uma descrição definida que seria satisfeita contingentemente por Édipo. Tal proposição seria o conteúdo (2) formulado acima:

(2) Édipo sente-se culpado.

Quer quando pense que ele próprio ou ele (essa pessoa designada ostensivamente) ou que Édipo está se sentindo culpado, Édipo estaria exprimindo um mesmo estado de coisas que tem a forma da proposição singular (2). À luz do princípio da simetria veritativa de Tugendhat, (4) não seria verdadeira a menos que (3) também o fosse. Entretanto, na medida em que Édipo pode acreditar que ele próprio está se sentindo culpado sem ter que acreditar que Édipo ou ele (indicando essa pessoa) estejam se sentindo culpados ou acreditar ainda que ele (indicando essa pessoa) ou que Édipo estejam se sentindo culpado sem ter que acreditar que ele próprio está se sentindo culpado, o conteúdo expresso em (2) (e o princípio da simetria veritativa) não são capazes de exprimir o que é próprio à auto-referência reflexiva.

A autoconsciência deve ser entendida como uma forma consciente de auto-referência em oposição a uma forma casual e não-consciente de auto-referência. Quando Édipo amaldiçoa o assassino de Laio, ele se refere a si mesmo sob a forma de um pensamento predicativo: o assassino de Laio está amaldiçoado. Mas ele não tem consciência que está se referindo a si mesmo uma vez que não sabe que a descrição definida “o assassino de Laio” se aplica unicamente a sua pessoa. Édipo só se refere conscientemente a si mesmo por meio de tal descrição definida quando, ao final da tragédia, reconhece que ele próprio é o assassino de Laio.

Essa capacidade peculiar de se referir conscientemente a si mesmo é caracterizada em orações principais <*oratio recta*> do discurso direto como (4) pelo empre-

go do dêitico essencial “eu” (Perry, 1979), e em orações subordinadas <*oratio obliqua*> em um relato sobre uma terceira pessoa pelo pronome reflexivo indireto “ele (ela) próprio” como um *quasi*-indicador (Castañeda, 1968), como no exemplo:

(13) Édipo pensa que *ele próprio\** está amaldiçoado<sup>1</sup>

A alternativa a uma teoria puramente referencial seria uma teoria descritiva do conteúdo da predicação de si capaz de dar expressão à reflexividade própria ao emprego do pronome de primeira pessoa. Segundo Reichenbach, só podemos entender como diferentes ocorrências <token> de expressões dêiticas e de demonstrativos de um mesmo tipo <type> poderiam fazer referência a coisas diferentes sob a suposição de que o emprego dessas ocorrências <tokens> faria referência a elas mesmas (reflexividade) (Reichenbach 1947). Assim, haveria uma sinonímia entre o pronome de primeira pessoa “eu” empregado em (4) e o misto de descrição e demonstração: “o locutor *dessa* frase em primeira pessoa” ou, alternativamente, “o pensador *desse* pensamento em primeira pessoa”. Assim, ao enunciar a oração predicativa (4) (ou nela pensar), Édipo estaria exprimindo o seguinte conteúdo:

(14) O pensador *desse pensamento em primeira pessoa* (4) sente-se culpado.

(14) nos permitiria compreender a diferença de significação cognitiva entre (4) e as demais asserções. Édipo pode acreditar que Édipo, ou que ele (essa pessoa), está se sentindo culpado sem ter acreditar que ele próprio\* está se sentindo culpado uma vez que as condições de verdade reflexivas de (4) são inteiramente distintas das condições de verdade reflexivas de (3) e (2). Enquanto (14) exprime as condições de verdade reflexivas de (4), as condições de verdade reflexivas de (2) e (3) teriam a forma das seguintes proposições, respectivamente:

(15) O indivíduo que responde pelo nome de “Édipo” em (2) sente-se culpado.

(16) O indivíduo que é referido pelo pronome “ele” em (3) sente-se culpado.

Segundo a teoria dística <token-reflexive>, a auto-referência reflexiva (4) seria

---

(1) Castañeda se utiliza do sinal <\*> para indicar o reflexivo indireto.

sempre mediada pelo próprio pensamento auto-referente reflexivo (14)<sup>2</sup>. Uma pessoa faria referência consciente a si mesma na medida em que empregasse uma predicação de si auto-referente. Essa abordagem se apresenta como uma conseqüência natural do truísmo de que o dêitico “eu” se refere a quem quer que o empregue. O emprego do pronome de primeira pessoa que reflete o domínio da sua semântica exige que o falante saiba que ele está se auto-referindo ao realizar tal emprego. (14) constitui, portanto, uma das coisas fundamentais que uma falante competente deve saber ao enunciar a predicação (4).

Mas a despeito das razões que militam a favor da teoria dítica, (14) e (4) não podem ser equivalentes uma vez que exprimem estados de coisas diversos. Quando Édipo enuncia ou pensa (4), ele está predicando da sua própria pessoa a propriedade de estar amaldiçoado. (4) é verdadeira quando o indivíduo Édipo satisfaz a propriedade de sentir-se culpado. Quando ele enuncia ou pensa (14), em contrapartida, ele está predicando a sensação de culpa de alguém que satisfaça a condição de ser o locutor ou o sujeito pensante de (4). (14) é verdadeira quando o locutor ou o sujeito pensante de (4) satisfaz a mesma propriedade de sentir-se culpado. Segundo a terminologia introduzida por Kaplan, enquanto (4) exprime uma *proposição singular* sobre Édipo, (14) exprime uma *proposição geral* que não é especificamente sobre Édipo, mas antes sobre uma condição de identificação que é satisfeita por Édipo: ser o locutor ou o sujeito pensante de (4) (Cf. Kaplan 1989).

Mas se a proposição singular (2) é incapaz de exprimir o que é próprio à autoconsciência, a proposição geral (14) parece nos conduzir de volta ao problema da circularidade levantado por Fichte. Segundo a teoria do pensamento de ordem superior (HOT) (que, como observamos na introdução desse trabalho, constitui o sucedâneo contemporâneo da tradicional teoria da reflexão), o indivíduo só poderia se auto-referir de forma consciente na medida em que a proposição em primeira pessoa (4) dispusesse a identificar a si mesmo como o autor de (4) ao enunciar ou

---

(2) “Teoria dítica” como tradução mais adequada de “token-reflexive theory” é uma sugestão do parecerista anônimo acolhida por mim.

entreter em pensamento a proposição de ordem superior expressa por (14) (conferir Rosenthal 2004, pág. 168). A esse respeito, Bermúdez observa o seguinte:

Parece claro que tal pessoa não terá êxito em referir-se a si mesmo a menos que ela saiba que ela produziu a ocorrência <token> em questão. Mas isso nos conduz diretamente de volta ao problema da circularidade. Empregar o pronome de primeira pessoa de modo a refletir o domínio da sua semântica (...) exige saber que se é o autor de tal emprego e isso é um conhecimento com um conteúdo de primeira pessoa. (Bermúdez, 1997 15).

Se, por um lado, é apenas por meio do emprego do pronome de primeira pessoa que consigo fazer referência a mim mesmo de forma consciente, por outro lado, não terei êxito em fazer referência a mim mesmo pelo emprego de tal pronome a menos que já saiba que sou o produtor dessa ocorrência. Assim compreendida, a dificuldade levantada originalmente por Fichte sobrevive ao giro lingüístico favorecido por Tugendhat. O problema levantado por Fichte repousa sobre a suposição geral de que a auto-referência reflexiva tem que ser ela própria um ato auto-consciente (no qual o sujeito que se auto-refere saiba que é ele quem realiza o ato de auto-referência). Assim, ou bem é necessário um conteúdo de ordem superior para que o sujeito possa se tornar consciente de que ele próprio está realizando tal ato, ou bem o sujeito se torna consciente de que está se auto-referindo no próprio ato original de auto-referência. No primeiro caso, nos vemos confrontados com um regresso infinito enquanto no segundo com um círculo vicioso.

Assim, o problema inicial da circularidade pode ser descrito semanticamente por meio de um dilema fundamental. Por um lado, a proposição singular (2) e o princípio da simetria veritativa são incapazes de exprimir o que é próprio à auto-referência cognitiva. Por outro lado, a proposição geral (14) parece pressupor a auto-referência cognitiva ao invés de explicá-la predicativamente.

#### REFERENCIALISMO CRÍTICO E TEORIA DA INFORMAÇÃO

Poderemos agora sobre a formulação semântica que Bermúdez apresenta para o problema da circularidade explicativa na estrutura da autoreferência refle-

xiva. O autor nos apresenta quatro formulações alternativas para aquela que seria a regra fundamental que regeria o emprego do pronome da primeira pessoa. Segundo a formulação inicial teríamos o seguinte:

Regra <token-reflexive rule> (primeira versão): ao empregar o pronome da primeira pessoa, a pessoa se refere a ela própria (Bermúdez, 1998, 14).

Na medida, entretanto, em que o falante (ou pensante) poderia estar se auto-referindo ao empregar o pronome de primeira pessoa sem o saber, essa formulação inicial seria incapaz de exprimir o emprego de tal pronome “de modo a refletir sua semântica”. A exclusão da possibilidade de tal auto-referência não-consciente nos leva a interpretar o pronome reflexivo “ela própria” como um *quasi*-indicador na acepção proposta por Castañeda. Assim, alcançamos uma segunda formulação:

Regra (segunda versão): ao empregar o pronome da primeira pessoa a pessoa se refere a ela\* própria (Bermúdez, 1998, 14).

A dificuldade óbvia dessa formulação reside no seu caráter imediatamente circular. Uma vez que o emprego do pronome reflexivo como um *quasi*-indicador só pode ser compreendido à luz do emprego do pronome da primeira pessoa, estaríamos explicando o emprego de tal pronome por meio de uma regra que só poderia ser compreendida quando por quem já houvesse compreendido a semântica do pronome da primeira pessoa. A alternativa, segundo Bermúdez, seria uma formulação sugerida Campbell (Cf. Campbell, 1994, 102) no seguintes termos:

Regra (terceira versão): qualquer emprego do “eu” faz referência a quem o emprega.

Entretanto, essa terceira variante da regra de emprego do pronome da primeira pessoa parece comportar duas interpretações. Segundo uma primeira leitura, essa terceira versão seria apenas uma variante notacional da primeira versão

(segundo a qual ao empregar o pronome da primeira pessoa, o falante se refere a ele próprio). Nela, o pronome reflexivo “ele próprio” não seria entendido como um *quasi*-indicador e deixaria em aberto a possibilidade de uma auto-referência não-reflexiva. Segundo Bermúdez, haveria, contudo, uma interpretação alternativa para a essa terceira versão que excluiria a possibilidade de uma auto-referência não-consciente sem recorrer ao quase-indicador “ele próprio” tal como a segunda versão:

Regra (quarta versão): Se uma pessoa emprega o pronome da primeira pessoa, então ela se refere a ela própria em virtude de ser o indivíduo que realiza tal emprego (Cf. Bermúdez, 1998, 15).

Essa versão exclui a possibilidade de que a pessoa que emprega o pronome de primeira pessoa desconheça que ela esteja se auto-referindo com o acréscimo da suposição de que ela sabe que está empregando tal pronome. Se essa (quarta) versão dítica <token-reflexive> estiver correta, a referência consciente que faço a mim mesmo por meio do emprego do pronome da primeira pessoa em (4) seria mediada pela referência reflexiva à ocorrência <token> do próprio pronome em (4) através de descrições demonstrativas da forma: o sujeito que realiza o emprego de tal pronome em (4). Mas aqui nos veríamos às voltas mais uma vez com o problema da circularidade. Como o saber que se é a pessoa que está empregando pronome da primeira pessoa constitui um conteúdo em primeira pessoa, estaríamos mais uma vez pressupondo um pensamento em primeira pessoa para compreendermos a regra que explicita o emprego do pronome da primeira pessoa.

A terceira versão da regra do emprego do pronome da primeira pessoa que Bermúdez encontra em Campbell constitui a bem da verdade uma variação de uma das três regras formuladas por Kaplan no seu célebre trabalho:

(D2) Em cada um dos seus proferimentos, o “eu” se refere a quem o profere (Kaplan, 1989a, 520).

Segundo o referencialismo de Kaplan, o pronome de primeira pessoa “eu” e a descrição demonstrativa “a pessoa que emprega esse pronome” não são expressões sinônimas. Se houvesse tal sinonímia, eu deveria deixar de existir quando deixasse de empregar tal pronome (Cf. 1989a, 520). Enquanto o conteúdo do pronome de primeira pessoa é o indivíduo (que o emprega), o conteúdo da descrição demonstrativa “a pessoa que emprega o pronome de primeira pessoa em (4)” é uma condição de identificação satisfeita por indivíduos. Assim, a referência cognitiva do pronome da primeira pessoa a quem o utiliza não é mediada por uma referência reflexiva a alguma ocorrência <token> desse mesmo pronome. Ela seria determinada diretamente pelo “caráter” do pronome da primeira pessoa, como uma função que associa a cada contexto de enunciação a pessoas que empregam do pronome da primeira pessoa nesses contextos. Isso por si só já exclui a interpretação favorecida por Bermúdez da terceira versão da regra nos termos da quarta versão (segundo a qual a referência consciente que o falante faria a si mesmo por meio do emprego do pronome da primeira pessoa seria mediada pela referência reflexiva a ocorrências <tokens> desse mesmo pronome sob a forma de descrições demonstrativas).

Mas o que leva Bermúdez a supor que a terceira versão da regra deveria ser interpretada nos termos da quarta versão é o reconhecimento da possibilidade de uma auto-referência não-reflexiva. Como observamos, uma auto-referência não-consciente ocorre toda vez que empregamos nomes próprios, descrições definidas e mesmo do demonstrativo “ele ali” sem que saibamos que somos os indivíduos a quem tais nomes, descrições ou demonstrativos se referem. Isso é possível porque o caráter de cada um desses termos singulares não estabelece qualquer conexão entre o contexto de enunciação de tais termos singulares e a referência dos mesmos. Entretanto, essa possibilidade está excluída no emprego do pronome da primeira pessoa. O caráter do pronome da primeira pessoa estabelece a conexão fundamental entre o contexto do seu emprego e o indivíduo referido por tal emprego em tal contexto. Como o próprio Bermúdez reconhece, um indivíduo não pode realizar um pensamento em primeira pessoa “sem saber que ele está pensando em si mesmo” (Cf. Bermúdez 1998, 3). Por conseguinte, o emprego do pronome de pri-



meira pessoa “de modo a refletir sua semântica” não exige a consciência prévia de que se é o indivíduo que está realizando tal emprego como afirma Bermúdez (Cf. Bermúdez, 1998, 15). A consciência de que se é a pessoa empregando o pronome da primeira pessoa não constitui um pressuposto, mas antes uma decorrência da consciência de si mesmo expressa por uma predicação de si em primeira pessoa. No nosso exemplo, Édipo pensa que ele\* próprio é a pessoa quem está empregando a predicação de si em primeira pessoa “sinto-me culpado” porque ele está pensando que ele\* (Édipo) está se sentindo culpado. Como supunha Tugendhat, a circularidade é um pseudo-problema.

Entretanto, mesmo que a auto-referência reflexiva não seja mediada pela referência reflexiva a uma ocorrência <token> do pronome da primeira pessoa por meio de descrições demonstrativas, o conteúdo reflexivo expresso por (14) desempenha um papel fundamental na compreensão da autoconsciência. Segundo o referencialismo crítico de Perry, sem o conteúdo reflexivo expresso por (14) não podemos compreender a *diferença cognitiva* fundamental entre a predicação de si em primeira pessoa (4) e as demais predicções de si com o mesmo conteúdo referencial (2) e (3). Édipo pode coerentemente pensar (4) e negar (2) ou, inversamente, pensar (2) e negar (4), porque, ao pensar (4), ele reconhece que ele próprio é o indivíduo está realizando o pensamento (4), o mesmo não ocorrendo ao pensar as demais predicções de si. Ao explicitar a conexão entre o contexto do emprego do pronome da primeira pessoa e a pessoa a quem tal pronome se refere, a proposição reflexiva (14) nos permite caracterizar o estado doxástico peculiar no qual o sujeito se encontra ao enunciar (4). No restante desse trabalho, buscarei uma explicação naturalista tanto para esse estado peculiar quanto para o conhecimento imediato de si que nele se baseia.

Desde o aparecimento de *The flow of Information* (Dretske 1981), a palavra “informação” adquiriu na filosofia um sentido técnico. Dretske caracteriza “informação” uma probabilidade condicional igual a  $Um$  de que uma situação  $S'$  ocorra quando outra  $S$  tiver ocorrido:  $Pr(S'/S) = 1$ . Segundo Perry e Barwise (1983), no entanto, há inúmeras relações de informação que não se deixam reduzir a probabilidades: relações necessárias no campo da matemática, leis naturais,

convenções lingüísticas etc. Assim, a noção de informação é melhor compreendida como uma relação entre tipos de situações relativas a uma restrição <constraint> (leis, regularidades nômicas etc.). Diremos então que uma situação  $S'$  (fato indicador) veicula informação sobre uma situação  $S$  (fato indicado) relativamente a uma restrição  $R$ :  $S's,R$ . Como informação não é uma propriedade, mas uma relação, um mesmo tipo de situação pode indicar diferentes tipos de situações relativamente a diferentes restrições. Assim, o mesmo fato de que anéis de uma árvore apresentam um determinado diâmetro podem indicar diferentes fatos ou situações relativamente a diferentes restrições: relativamente a uma determinada restrição, indica a idade da árvore ( $R1$ : se os anéis de uma árvore apresentam um determinado diâmetro, então a árvore possui determinada idade); relativamente a outra restrição, indica a população de formigas ( $R2$ : se os anéis apresentam tal diâmetro, então há uma quantidade determinada de formigas que nela habitam); e relativamente a uma terceira restrição, indica a quantidade de chuva a que a árvore está submetida ( $R3$ : se os anéis apresentam diâmetro, então o índice pluviométrico na região é tal e qual).

Entretanto, a teoria da informação não pode ser tomada ainda como uma teoria da referência ou da representação (como Dretske supunha em 1981). Enquanto a referência sempre supõe a possibilidade de uma referência equivocada <misrepresentation>, não faz sentido falarmos em informação equivocada <misinformation> na acepção técnica do termo. O estatuto nômico da restrição relativamente a qual os tipos de situação se conectam exclui a possibilidade de que uma situação deixe de indicar a outra. O fato de que a árvore apresenta um determinado diâmetro não pode deixar de indicar o fato de que ela possui uma determinada idade segundo uma determinada restrição ( $R1$ ), como não pode deixar de indicar que o índice pluviométrico na região segundo outra restrição ( $R3$ ). Tendo tal dificuldade em mente, Dretske passa a advogar nos seus dois últimos trabalhos (Dretske 1988, 1995) uma abordagem mista da noção de representação, combinando a teoria da informação com um elemento teleológico. Um tipo de situação representa  $S'$  outro tipo de situação  $S$  não apenas quando há uma lei conectando  $S'$  e  $S$ .  $S'$  só representa  $S$  quando adquire a *função* de indicar  $S$ . Assim, relativamente a uma determinada

restrição, a percepção do vermelho não se limita a indicar uma determinada propriedade (vermelho) do corpo que estimula a visão. Ela também *representa* tal propriedade quando o sistema cognitivo adquire a *função* de indicar tal propriedade, o que ocorreria ou por evolução natural ou por aprendizagem do predicado “vermelho”. No primeiro caso teríamos o que Dretske denomina indicador de função sistêmico, enquanto no segundo um indicador de função adquirido.

A distinção usual entre representações sensíveis e conceituais passa a ser compreendida nos termos da distinção entre esses dois indicadores de função. Uma representação é de natureza não-conceitual quando seu indicador de função é derivado do sistema cognitivo do qual a representação é um estado que guarda relações para com outros estados do mesmo sistema cognitivo. Neste caso, o indicador de função é adquirido filogeneticamente pela seleção natural e a restrição assume a forma de uma regularidade biológica. Por outro lado, uma representação é de natureza conceitual quando o indicador de função é adquirido pelo aprendizado de que a situação na qual o sistema se encontra é uma ocorrência <token> de um tipo <type> de situação (fato indicador) que envolve outro tipo de situação (fato representado). Nesse caso, a restrição assume a forma de uma regra semântica.

Isso nos permite caracterizar o estado doxástico de quem pensa um conteúdo de si em primeira pessoa nos seguintes termos. Em primeiro lugar, o aprendizado do predicado envolvido confere ao estado a função de indicar que um tipo determinado de propriedade (a propriedade de sentir-se culpado) está sendo instanciada. E o aprendizado do caráter do pronome de primeira pessoa (como uma função que leva do contexto do seu emprego à pessoa quem o emprega) confere ao estado em questão a função de indicar que quem quer que se encontre em tal estado (Édipo) está instanciando o tipo de propriedade em questão (sentir-se culpado). Mas como o caráter do pronome em primeira pessoa conecta o contexto do seu emprego com a pessoa quem o emprega (Édipo), esse estado também veicula a informação reflexiva que o indivíduo que instancia a propriedade (sentir-se culpado) é o mesmo que entretém tal pensamento em primeira pessoa “sinto-me culpado”.

É nesse plano da informação reflexiva que emerge a diferença cognitiva entre o estado doxástico de quem entretém uma predicação de si em primeira pessoa (4) e os estados doxásticos de quem entretém predicacões de si que empregam nomes próprios (2) ou o demonstrativo “ele ali” (3) como termos singulares. Tendo aprendido que o caráter do nome próprio “Édipo” é uma função constante cujo valor é sempre a pessoa de Édipo em qualquer contexto, o estado no qual a pensa se encontra ao entreter (2) veicula a informação reflexiva de que o indivíduo que responde pelo nome de “Édipo” em (2) se sente culpado. Tendo também aprendido que o caráter do demonstrativo “ele ali” é uma função variável cujos argumentos são os contextos de demonstração e os valores os indivíduos indicados, ao entreter (3) o sujeito se encontra em um estado mental que veicula a informação reflexiva de que o indivíduo que é referido pelo demonstrativo “ele ali” em (3) se sente culpado. Mas tendo aprendido o caráter do pronome da primeira pessoa como a função que tem por argumento o contexto do emprego de tal pronome e o valor a pessoa quem o emprega, o estado de quem entretém o pensamento em primeira pessoa (4) veicula a informação reflexiva de que o próprio sujeito pensante de (4) se sente culpado. Uma vez que a predicação de si em primeira pessoa (4) apresenta o mesmo conteúdo referencial que as demais (2) e (3) (a proposição singular constituída pelo indivíduo Édipo e a propriedade de sentir-se culpado) apenas a informação reflexiva veiculada por (4) é capaz de explicar porque Édipo se cega e se condena ao exílio após entreter (4).

Mas só podemos falar de conhecimento em sentido próprio, quando esse estado doxástico cumprir sua função de indicativa, ou seja, veicular a informação substantiva de que o sujeito que nele se encontra (Édipo) está instanciando a propriedade em questão (sentir-se culpado). A satisfação da sua função adquirida pode ocorrer de duas formas. Segundo uma primeira possibilidade, o estado doxástico de Édipo ao entreter (4) estaria sendo causado por estados perceptivos que estariam cumprindo a sua função de veicular a informação substantiva de que Édipo é a pessoa quem estaria instanciando a propriedade de sentir-se culpado. O estado de Édipo ao entreter o pensamento em primeira pessoa (4) poderia estar sendo causado ou por ponderações de uma terceira pessoa (um psicanalista que

tenha observado seu comportamento) ou pela observação que ele próprio poderia fazer da sua conduta.

De acordo com uma segunda possibilidade, contudo, o estado doxástico de Édipo ao entreter (4) poderia ser causado por um estado sensorial não-perceptual, cumprindo sua função de indicar que ele\* próprio estaria se sentindo culpado. Édipo saberia de forma imediata, portanto, que ele\* próprio se sente culpado quando seu estado doxástico ao entreter (4) resultasse apenas da sua sensação de culpa. Assim, como supunha Evans (1982), a ausência de identificação não constitui uma propriedade do conteúdo expresso por orações psicológicas em primeira pessoa, mas antes das bases epistemológicas do estado de quem entretém o pensamento (4).

Mas, como salientamos, nem sempre que alguém se auto-atribui uma condição física ou mental é fato que se encontra em tal condição ou é fato que é a pessoa em questão que se encontra na condição em questão (sob a suposição de que alguém se encontra na condição em questão). O erro ocorre quando o estado doxástico no qual Édipo se encontra, tendo adquirido a função de indicar que Édipo está se sentindo culpado, é causado por alguma outra propriedade ou por alguém que não o próprio Édipo, segundo conexões nomológicas distintas. No âmbito do auto-conhecimento imediato, há duas formas possíveis de erro. Em primeiro lugar, o indivíduo poderia incorrer em um erro de conceituação acerca da propriedade que ele estaria instanciando. Assim, Édipo poderia crer erroneamente que ele\* está sentido culpa (quando a bem da verdade sente, digamos, vergonha) quando ele não soubesse exatamente o que significa o predicado “sentir-se culpado”.

Na outra situação possível, o pensamento (4) de Édipo poderia se revelar falso porque suas sensações estão representando erroneamente (de forma não-conceitual) o estado no qual ele se encontra. À luz da teoria da informação, podemos compreender a familiaridade consigo mesmo <Kenntnis seiner selbst> como a informação sistêmica veiculada por uma série de estados cerebrais conectados nomicamente a propriedades do sistema cognitivo. Além das sensações usuais de prazer e desprazer, essa forma primitiva não-conceitual de familiaridade consigo mesmo envolve inúmeros sensores especializados em uma série de sistemas infor-

macionais distintos e independentes, que não cabe aqui especificar em detalhe. Em linhas gerais, determinados sistemas têm a função de fornecer informação sobre o interior do corpo, enquanto outros a função têm de prover a informação sobre a orientação e movimento do corpo como um todo (sistema vestibular) ou dos seus membros (sistema proprioceptivo). Quando tais estados são causados por propriedades ou por outros indivíduos que eles não têm por função sistêmica indicar, eles passam a representar erroneamente ou as propriedades ou o próprio indivíduo as estaria instanciando.

Como os estados doxásticos de quem entretém um pensamento em primeira pessoa se originam em tais estados sensoriais não-conceituais, um erro no âmbito não-conceitual acaba por acarretar um erro judicativo. Édipo acredita erroneamente que sente culpa quando sua sensação representa erroneamente seu estado como culpa. E acredita erroneamente que é ele\* quem sente culpa quando seu pensamento (4) for causado pela sensação real de culpa de uma terceira pessoa. Contudo, ainda que tal estado doxástico de Édipo ao entreter o pensamento em primeira pessoa (4) não veicule a informação substantiva de que próprio Édipo está se sentindo culpado, ele nunca deixa de veicular a informação reflexiva de que o próprio sujeito pensante de (4) está sendo *representado* em (4) como aquele que se sente culpado.

#### RESUMO

*O artigo consiste em uma leitura crítica da abordagem de Tugendhat da auto-consciência à luz da filosofia contemporânea da mente. Rejeitando a abordagem "continental", baseada no tradicional modelo sujeito-objeto, Tugendhat fornece uma abordagem alternativa da auto-consciência baseada no que ele chama de "abordagem analítica", de acordo com a qual a auto-consciência exibe a forma de um pensamento-do-eu (Ich-sätze) psicológico. Neste artigo, após apresentar a posição de Tugendhat, argumento em favor de uma abordagem informacional da auto-referência cognitiva e do auto-conhecimento imediato como solução para os enigmas tradicionais acerca da auto-consciência.*

**Palavras-Chave:** Auto-consciência, Tugendhat.

ABSTRACT

*This paper is a critical reading of Tugendhat's account of self-consciousness in the light of contemporary philosophy of mind. Rejecting the continental approach, based on the traditional subject-object model, Tugendhat provides an alternative account of self-consciousness based in what he calls "language-analytical approach", according to which self-consciousness exhibits the form of a psychological I-Thought <Ich-Sätze>. In this paper, after assessing Tugendhat's position, I shall advocate an information-based account of cognitive self-reference and immediate self-knowledge as the solution for the traditional puzzles about self-consciousness.*

**Keywords:** self-consciousness; Tugendhat.

Recebido em 11/2006

Aprovado em 09/2007

*Referências:*

Bermúdez, José. 1998. *The paradox of self-consciousness*. Cambridge: MIT Press.

Castañeda, H. N., 1966. "He": A study in the logic of self-consciousness. *Ratio*, 8: 130-57.

Castañeda, H. N. "Indicators and Quasi-Indicators". *American Philosophical Quarterly*, 4, 1967: 85--100.

Castañeda, H. N. 1989. "The reflexivity of self-consciousness: Sameness/identity, data for artificial intelligence". *Philosophical Topics*, 17:27-58.

Cramer, K. 1970. *Erlebnis*. Em: *H. Gadamer , Stuttgarter Hegel Tage*. Bonn.

Dretske, F., 1981. *The Flow of information*. Bradford Books, MIT Press, Cambridge, Massachusetts.

Dretske, F., 1995. *Naturalizing the mind*. MIT Press.



Evans, G, 1982, *Varieties of Reference*, Oxford University Press.

Fichte, J.G., 1794. *Grundlage der gesamten Wissenschaftslehre*. Jena und Leipzig.

Henrich, D. 1967. *Fichtes ursprüngliches Einsicht*. Frankfurt a. M.

Henrich, D.1970. "Selbstbewusstsein: kritische Einleitung in eine Theorie". Em: R. Bubner (Editor), *Hermeneutik und Dialekt*, Tübingen.

Kaplan, David. "Afterthoughts". Em (Almog, et. al., 1989a), 565-614.

Kaplan, David. "Demonstratives". Em (Almog, et. al., 1989b), 481-563. (Manuscrito original de 1977).

Perry, John. 1979. "The Problem of the Essential Indexical". *Noûs* 13: 3-21.

Perry, John, Barwise, Jon. 1983. *Situations and Attitudes*. Cambridge, MIT Press.

Perry, John. 1985. "Self-Knowledge and Self-Representation," *Proceedings of IJCAI* (1985): 238-42.

Perry, John. 1993. *The Problem of the Essential Indexical and Other Essays*. New York: Oxford University Press.

Perry, John. 2001. *Reference and Reflexivity*. Stanford: CSLI Publications.

Pothast, U. 1971. *Über einige Fragen der Selbstbeziehung*. Frankfurt am Main: Vittorio Klostermann.

Reichenbach, Hans. "Token-reflexive words". Em: *Elements of Symbolic Logic*. New York: The Free Press, 1947: 284ff.

Rosenthal, D., 2004. "Being conscious of ourselves". *Monist* 87, 2.

Shoemaker, Sydney. 1963. *Self-knowledge and self-identity*. Ithaca, NY: Cornell University Press.

Shoemaker, Sidney. 1968. "Self-Reference and Self-Awareness". *Journal of Philosophy* 65: 555-567.

Tugendhat, E. 1979. *Selbstbewusstsein und Selbstbestimmung. Sprachanalytische Interpretationen*. Tradução inglesa: P. Stern. *Self-Consciousness and Self-Determination*. Cambridge, MA: MIT Press, 1986.

Tugendhat, E. 1997. *Egozentrität und Mystik*. München.1999. Citado a partir da tradução espanhola de Maricio Suárez Crothers *Egocentricidad y Mística*. Barcelona 2004.

Wittgenstein, 1958. *The Blue and Brown Books*. Oxford: Blackwell.