

Editorial

This issue of *Disputatio* is dedicated to the memory of Manuel dos Santos Lourenço (1936–2009). M. S. Lourenço introduced analytic philosophy in Portugal and taught most of the present members of the editorial committee of *Disputatio*. It can safely be said that he was the first causal link in the creation of a national analytic philosophy community — without M. S. Lourenço, we probably would not have been here as members of a community of Portuguese analytic philosophers. He was also the first editor of *Disputatio* from 1996 to 2004. The editors thus pay him a deserved homage.

M.S. Lourenço studied under the supervision of Michal Dummett at the University of Oxford (M.A.) and prepared then *O Teorema de Gödel e a Hipótese do Contínuo* (F. C. Gulbenkian, 1979) for publication. A second edition came out in 2009. He also translated Wittgenstein's *Tractatus* and *Investigations* into Portuguese (*Tratado Lógico-Filosófico e Investigações Filosóficas*, F. C. Gulbenkian, 1987). He completed his PhD in 1980 at the University of Lisbon, with the dissertation *Espontaneidade da Razão: A Analítica Conceptual da Refutação do Empirismo na Filosofia de Wittgenstein* (which was published as a book in 1986 by Imprensa Nacional-Casa da Moeda). He also published *Teoria Clássica da Dedução* (Assírio e Alvim, 1991), and the book *A Cultura da Subtileza: Aspectos da Filosofia Analítica* (Gradiva, 1995, edited by Desidério Murcho). The latter collected radio dialogues with several philosophers and other Portuguese intellectuals in the programme *Rádio Cultura*, transmitted in RDP2.

We publish now a small sample of his published books and articles, and some of his unpublished writings. The first consists of a selection of short introductions to different subjects originally published in *A Cultura da Subtileza*, and attests Lourenço's wide variety of interests. The second article, 'Um Filósofo da Evidência,' was first published in *Boletim da Sociedade Portuguesa de Matemática*, (2006). The third, 'The Axioms of Wisdom' is a short original note. 'What is Structuralism?' is the text of the inaugural lecture given by M.S. Lourenço in the first meeting of the Portuguese Society of Analytic Philosophy, which took place in Coimbra in 2002. Finally, 'A Refutação da Teoria Empirista da Consciência' is the second chapter of *A*

Espontaneidade da Razão (presently out of print and hence inaccessible to new readers).

We thank Nuno Jerónimo and Ana Cristina Domingues for making available some of the original material here included, and also José Martínez, Rodrigo Alexandre de Figueiredo e Delvair Custódio Moreira. We are especially grateful to the descendants of M.S. Lourenço, particularly to Frederico Lourenço, for the generous authorization to use the texts that we now publish.

The editors

Editorial

Este número da *Disputatio* é dedicado à memória de Manuel dos Santos Lourenço (1936–2009), que introduziu a filosofia analítica no nosso país, consolidou o ensino da lógica formal da Faculdade de Letras da Universidade de Lisboa e foi professor da maioria dos membros da direcção da *Disputatio*. Podemos dizer com segurança que foi o primeiro elo causal na criação da comunidade analítica nacional — sem ele, muito provavelmente não estaríamos aqui enquanto membros de uma comunidade de filósofos analíticos portugueses. Foi ainda o primeiro editor da revista *Disputatio*, de 1996 a 2004. A direcção da *Disputatio* presta-lhe assim esta merecida homenagem.

M.S. Lourenço estudou sob a orientação de Michael Dummett na Universidade de Oxford (M.A.) e preparou nessa altura *O Teorema de Gödel e a Hipótese do Contínuo* (Gulbenkian, 1979), que teve uma segunda edição em 2009. Traduziu ainda o *Tractatus* e as *Investigações* de Wittgenstein (Gulbenkian, 1987). Em 1980, doutorou-se na Universidade de Lisboa com a dissertação *Espontaneidade da Razão: A Analítica Conceptual da Refutação do Empirismo na Filosofia de Wittgenstein* (que foi publicada em livro em 1986 pela Imprensa Nacional Casa da Moeda). Publicou ainda *Teoria Clássica da Dedução* (Assírio e Alvim, 1991), e o livro *A Cultura da Subtileza: Aspectos da Filosofia Analítica* (Gradiva, 1995). Este último reuniu diálogos com várias personalidades da cultura portuguesa no programa Rádio Cultura, da RDP2.

Publicamos agora alguns extractos de livros ou textos de M. S. Lourenço, incluindo alguns inéditos. O primeiro consiste numa selecção de pequenas introduções a diferentes temas originalmente publicadas em *A Cultura da Subtileza*, que demonstram a diversidade

dos seus interesses filosóficos. O segundo, «Um Filósofo da Evidência», foi originalmente publicado no *Boletim da Sociedade Portuguesa de Matemática* (2006). O terceiro, «The Axioms of Wisdom» é uma pequena nota inédita. «What is Structuralism» foi o texto apresentado por Manuel Lourenço na palestra inaugural do primeiro Encontro Nacional de Filosofia Analítica, que decorreu em Coimbra em 2002. Finalmente, «A Refutação da Teoria Empirista da consciência» é o segundo capítulo de *A Espontaneidade da Razão* (actualmente esgotado e portanto inacessível aos leitores contemporâneos).

Agradecemos a Nuno Jerónimo e a Ana Cristina Domingues por nos disponibilizar algum do material aqui incluído, assim como a José Martínez, Rodrigo Alexandre de Figueiredo e Delvair Custódio Moreira. Agradecemos principalmente aos descendentes de M. S. Lourenço, em particular a Frederico Lourenço, pela generosa autorização a utilizar os textos que agora publicamos.

Os editores

Da cultura da subtileza*

M. S. Lourenço

Universidade de Lisboa

Um facto crucial no desenvolvimento da filosofia do século que agora termina foi a fragmentação da cultura filosófica europeia no que na verdade se pode considerar duas culturas filosóficas separadas, como se reflecte imediatamente na análise da literatura produzida neste século: há manifestamente uma cultura filosófica das línguas românicas (ou dos países destas línguas) cuja teoria e prática é irreconciliável com a cultura filosófica que partindo do Círculo de Viena se veio a generalizar aos países anglo-saxónicos e hoje a uma parte considerável da produção contemporânea de língua alemã. Sem querer entrar agora numa caracterização pormenorizada das diferenças que separam as duas culturas, gostaria no entanto de escolher dois critérios que facilmente isolam os aspectos vivamente antagónicos das duas culturas filosóficas, um primeiro critério de carácter cognitivo e um segundo de carácter estético ou estilístico. Assim, do ponto de vista cognitivo, a filosofia que é geralmente associada com a literatura de línguas românicas prossegue um objectivo que se pode, à falta de melhor, designar por especulativo, no sentido em que em geral se pretende à custa da associação de certas palavras ou grupos de palavras *evocar* pensamentos e conceitos, não sendo no entanto de todo possível provar que tal evocação é univocamente realizada pelo grupo de palavras escolhido para a representar. Em contraste, a cultura filosófica que hoje se implantou nos países anglo-saxónicos e num número crescente de universidades de língua alemã, conhecida através da designação de filosofia analítica, tem uma finalidade estritamente cognitiva, no sentido em que os seus cultores procuram através de um argumento válido demonstrar um aspecto até agora irreconhecido da estrutura do nosso conhecimento, sujeitando-se para isso a um veredicto sobre se a sua pretensão cognitiva é de aceitar ou de recusar. Os

* Segue-se cinco textos diversificados e um extracto do Prefácio de *A Cultura da Subtileza*, de M. S. Lourenço (Lisboa: Gradiva, 1995).

clássicos deste modo de escrever filosofia erigiram em instrumento decisivo para o refinamento dos seus argumentos o *culto da subtileza*, a partir do qual também se torna possível avaliar uma teoria filosófica como interessante por ser subtil ou trivial por não ter o grau suficiente de subtileza que permita distinções conceptuais diferenciadas. E assim, passando agora para o ponto de vista estético ou estilístico, esta cultura filosófica tem que ser expressa numa prosa de recorte conceptista e não é por isso de admirar que Wittgenstein seja ao mesmo tempo um paradigma desta cultura e um artista da prosa de língua alemã deste século. Quanto à caracterização estilística da filosofia que é em geral produzida nos países de línguas românicas, ela parece ser uma reencarnação do estilo de Villemain, sobre o qual Baudelaire fez a análise definitiva no seu ensaio *L'Esprit et le Style de M. Villemain*:

Horreur congéniale de la clarté, dont le signe visible est son amour du style allusionnel [...] La phrase de Villemain, comme celle de tous les bavards qui ne pensent pas [...] commence par une chose, continue par plusieurs autres, et finit par une qui n'a plus de rapporte avec les précédentes que celle-ci entre elles.

O meu subtítulo «Aspectos da Filosofia Analítica» foi escolhido para sublinhar o carácter incompleto da apresentação feita da teoria e da prática da filosofia analítica, em particular nas suas relações com as subdisciplinas tradicionais da filosofia, uma incompletude que tem uma origem puramente contingente ou circunstancial. Das subdisciplinas tradicionais falta acima de tudo uma apresentação do que é o tipo de trabalho da filosofia analítica na ética, sobre o qual existe uma literatura copiosíssima e da qual se pode dizer que os seus resultados têm um grau de originalidade — e assim de subtileza — igual ao que foi atingido na filosofia da linguagem. Falta a seguir a apresentação da nova subdisciplina filosófica que se pode com direito dizer que foi uma criação da filosofia analítica em geral e de Wittgenstein em particular, nomeadamente aquilo a que ele chama filosofia da consciência e que foi fatal e erroneamente traduzida em inglês como *philosophy of mind*. Por último falta um relato do que tem sido o progresso feito em áreas da filosofia aplicada, como a filosofia do direito, a filosofia política e a filosofia da ciência em geral, uma designação que hoje exclui a filosofia da matemática. Chego assim propriamente ao momento da *captatio benevolentiae*, por meio do qual procuro arguir que a escolha apresentada é apesar de incompleta no entanto

suficientemente representativa do espírito e da letra da filosofia analítica de modo a merecer um estudo, por parte do leitor interessado, do que diz respeito às subdisciplinas da lógica, teoria do conhecimento, estética e filosofia da linguagem.

1. Filosofia e estética da música

Embora as expressões «filosofia da música» e «estética da música» pareçam muitas vezes ocorrer como sinónimas, quer na linguagem corrente quer na literatura filosófica, é necessário e conveniente separar os dois conceitos, de modo a atribuir à filosofia da música o papel da construção de conjecturas acerca da natureza última da música, e à estética da música o papel de uma análise dos problemas colocados pela análise filosófica da obra de arte musical. Assim, enquanto as conjecturas avançadas pelos filósofos da música não se deixam avaliar quanto à sua correcção, as soluções propostas para os problemas estéticos suscitados pela obra de arte musical são soluções racionais (ou pelo menos racionalizáveis) e deve ser possível proceder à sua avaliação.

Enquanto a lista dos filósofos que se têm ocupado com a determinação especulativa da natureza última da música é impressionante quanto ao seu comprimento, e inclui quase todos os nomes de vulto desde Platão a Husserl, os filósofos do século XIX têm tido no entanto um eco maior do que todos os outros. Esta tendência foi certamente iniciada por Schopenhauer o qual escreveu acerca da música num livro em princípio fora do domínio da estética, e definiu a essência do pensamento musical como sendo uma expressão directa da vontade. Particularmente lisonjeiro para os filósofos foi a sua doutrina, associada àquela, de que uma tal expressão directa da vontade é capaz de representar a essência metafísica do mundo, de tal modo que a música é ela própria uma metafísica sem palavras, e assim Schopenhauer chega à conclusão anti-aristotélica de que a música é a filosofia na sua forma verdadeira. Ele foi ao ponto de construir um pequeno aforismo para a sua tese, que é uma paródia de um aforismo mais conhecido de Leibniz sobre as relações inconscientes entre o pensamento musical e o pensamento aritmético, e ao qual Schopenhauer deu a seguinte forma: «A música é um exercício inconsciente de metafísica, no qual a alma não sabe que está a fazer filosofia.» No século XIX desempenhou ainda também um papel importante a filosofia da música de Nietzs-

che, o qual era ele próprio um músico amador, com a sua ideia de que esta é a arte dionisíaca *katexokén*, e ambas as filosofias serviram para a formação da *Weltanschauung* e da filosofia da música que se encontra na ópera e nos ensaios de Wagner, e nestes a sua dívida para com Schopenhauer é explicitamente mencionada.

Em geral pode-se dizer que os problemas da estética da música são mais acessíveis ou mais tratáveis do que as conjecturas propriamente especulativas da filosofia da música, e entre aqueles problemas há que fazer ainda uma distinção entre os que a música tem em comum com as outras artes daqueles que são específicos da obra de arte musical. Da primeira classe de problemas, daqueles que a música tem em comum com as outras artes, avulta o problema básico da representação da realidade na obra de arte musical. A primeira solução para este problema foi a de Aristóteles, segundo o qual a relação de representação é concebida como uma imitação ou mimese do objecto pela obra de arte. Sem dúvida que um proponente convicto da doutrina aristotélica consegue encontrar diversas instâncias de mimese propriamente musical, desde o galope dos cavalos de Monteverdi ao canto do rouxinol e do cuco na *Sinfonia Pastoral* de Beethoven ou ao ritmo do ferreiro na ópera *Siegfried* de Wagner: mas em breve os exemplos se esgotam e a generalidade das asserções musicais fica por esclarecer, se nos limitamos apenas ao domínio da doutrina da imitação.

Nestas circunstâncias o fundador da moderna estética da música como disciplina filosófica, Eduard Hanslick, formulou em 1854 a primeira teoria lógico-filosófica da obra de arte musical na sua obra *The Beautiful in Music*. Nesta obra Hanslick considera pela primeira vez as consequências que advêm para a teoria estética quando se interpreta a música como uma linguagem, para a seguir poder discutir o problema da denotação dos símbolos desta linguagem. A sua doutrina era a de que os símbolos da linguagem musical não têm denotação, e assim as únicas relações que se podem estabelecer entre eles são relações a que hoje chamaríamos sintácticas ou formais, razão pela qual a sua doutrina é justamente conhecida como formalismo musical. É disputável se no texto de Hanslick se pode acomodar ainda um esboço de semântica ou sentido para os símbolos musicais, ou se se está definitivamente perante um formalismo radical.

Por razões que não pertence agora enumerar sou a favor da primeira interpretação, sobretudo por ela proporcionar também uma base para a solução de um problema estético específico da música, que é a definição do estatuto ontológico da obra de arte musical. Uma

formulação simples deste problema é a seguinte: partindo de um ponto de vista extensional, o estatuto ontológico da obra de arte é em geral bem definido no interior de uma ontologia de objectos, uma vez que o objecto quadro é a obra de arte pictórica, o objecto catedral a obra de arte arquitectónica, etc. O problema transforma-se em crise quando se tenta manter o ponto de vista extensional, e portanto uma ontologia de objectos, em relação à obra de arte musical. Para esta, os únicos candidatos ao estatuto de objecto são a partitura e a realização física ou sonora desta. Mas ninguém está realmente disposto a aceitar que a obra de arte musical é o conjunto das notas impressas, uma vez que estas não constituem uma experiência acústica. Mas que a obra de arte seja a realização física da partitura está sujeita à dificuldade genérica da multiplicidade de realizações possíveis. Assim a determinação do estatuto ontológico da obra de arte musical parece ser incompatível com o princípio da extensionalidade.

2. Filosofia e introspecção cognitiva

O que torna a nossa experiência mental surpreendente é ela ser incompreensível, outras vezes é ser ambivalente, outras vezes ainda é ser paradoxal. E assim o ponto de vista filosófico por excelência é estar surpreendido, ou estar perplexo, ou estar perante uma contradição. Em todo o caso, como a capacidade de se deixar surpreender é multiforme ou polimorfa, torna-se necessário ter uma ideia da direcção que a surpresa filosófica toma, a fim de esboçar nitidamente os seus contornos. O objectivo ou o fim da minha perplexidade filosófica é determinado pela minha vontade de eventualmente vir a ser capaz de distinguir aquilo que me aparece como sendo uma ilusão, daquilo que quero considerar ser conhecimento. Assim, o impulso que me leva a fazer perguntas, a descompreender, a querer compreender, é um impulso de conservação da minha experiência mental com vista a excluir na medida do possível a sua falsificação, não pelo mundo exterior, mas pela minha incapacidade de me separar de uma ilusão.

A filosofia tem assim um objectivo a que poderia chamar «sublime», e que consiste em distinguir o conhecimento da ilusão e, nestes termos, torna-se necessário postular a existência de uma faculdade mental capaz de fazer essa discriminação. Vou dar a essa faculdade o nome de «faculdade de introspecção cognitiva», a qual difere da introspecção usual por não ser um mero registo de dados da consciên-

cia, mas antes a descoberta do seu significado cognitivo. O resultado do uso desta faculdade é assim um acto cognitivo, representável afinal numa expressão da nossa linguagem. O fim da introspecção cognitiva não é a produção de novos dados, mas antes a descoberta de relações interessantes entre os dados já disponíveis.

Embora me tenha exprimido sempre como se estivesse apenas no domínio da realidade psíquica, fora de qualquer juízo de valor, e falasse da introspecção cognitiva como sendo uma função neutral, como a audição ou a visão, devo dizer imediatamente que a teoria da introspecção cognitiva tem implícita um juízo de valor acerca do resultado do uso dessa faculdade, como se vê pelo facto de eu ter isolado apenas aquelas relações supostas serem interessantes.

E assim, apesar da introspecção cognitiva poder levar à descoberta de um facto, os partidários da teoria da introspecção cognitiva não gostariam no entanto de dizer que a descoberta de qualquer facto implica o uso dessa faculdade. Do mesmo modo, embora a introspecção cognitiva conduza à verdade, nem todas as verdades merecem ser consideradas o produto do uso da introspecção cognitiva. Por exemplo, é um facto e uma verdade que todos os círculos são redondos, mas não é necessário qualquer poder de introspecção cognitiva para o descobrir. Assim, o teste crucial acaba por ter lugar quando as relações interessantes, descobertas entre um conjunto de dados, deixaram de estar inconscientes e passaram a conscientes, deixaram de estar simplesmente latentes e passaram a manifestas: assim, a introspecção cognitiva é a capacidade de tornar consciente o que até aí era inconsciente e por isso desconhecido.

A palavra «resultado» é usada no vocabulário científico como sendo equivalente a um facto científico ou uma verdade, e nesse sentido não é sujeito a revisões sucessivas, tem um efeito cumulativo, no sentido em que novas verdades podem ser obtidas a partir destas e em geral obrigam, uma vez formuladas, toda a comunidade científica. Neste sentido da palavra, a filosofia não apresenta resultados. As teorias filosóficas estão constantemente sujeitas a revisão, não têm qualquer poder cumulativo e estão longe de obrigar a comunidade filosófica. Fica-se assim diante da questão de saber como é que a filosofia pode sobreviver à sua incapacidade de produzir resultados. É a filosofia de todo incapaz de produzir resultados?

3. Cinema, guião e jogos de linguagem

Na percepção do filme de arte é ao fim de pouco tempo óbvio que se está diante de pelo menos duas obras de arte, a obra de arte visual constituída pela encenação e a obra de arte acústica ou literária constituída pelo guião. Nestas circunstâncias o filme de arte proporciona um *analogon* da experiência estética da ópera, onde também em princípio, embora nem sempre *actualiter*, pelo menos duas e no máximo três artes actuam simultaneamente. Esta era de resto a ideia subjacente à teoria wagneriana da ópera como *Gesamtkunstwerk* e é neste sentido que o filme de arte é uma realização da teoria do efeito de síntese produzido pelas artes ao realizarem a concretização da intenção poética.

De acordo com esta teoria estamos em geral perante duas situações básicas: ou o efeito de síntese é conseguido e não há lugar a uma avaliação hierárquica, ou o efeito de síntese falha e é-se forçado a distinguir valorativamente as diferentes contribuições das artes componentes para o efeito total.

Mas se esta teoria for comparada com os factos da experiência ao nosso dispor salta em primeiro lugar à vista que as situações que se encontram nas óperas *Electra* e *Salomé*, em que duas obras de arte literárias dão origem a duas obras de arte musical, constituem uma amostra irrepresentativa, e não faria por isso sentido em só a seu respeito falar de um efeito de síntese.

Em segundo lugar, o wagneriano mais fiel não consegue iludir o facto de que o *libretto* wagneriano não constitui uma obra de arte literária, caso em que se é forçado a reconhecer que a ópera de Wagner não consegue justamente atingir o efeito de síntese que era suposta ser capaz de exemplificar.

Nestas circunstâncias, o princípio de síntese idealizado na teoria da *Gesamtkunstwerk* não tem uma justificação na experiência, uma vez que enquanto em teoria as artes componentes da obra de arte total estariam subsumidas no sentido da convergência e não haveria assim lugar à definição de uma hierarquia entre elas, na prática esta convergência dá frequentemente lugar à divergência, e logo à possibilidade de as separar e avaliar hierarquicamente.

É assim surpreendente que este injustificado princípio de síntese continue a modelar os nossos hábitos mentais e a ser um dos obstáculos — mas não o único — a considerar como obras de arte autónomas algumas das artes componentes, quer na ópera, quer no filme de arte.

É interessante por isso analisar o género de dificuldades teóricas que militam contra o reconhecimento do guião do filme de arte como obra de arte literária, um problema que é cognato dos problemas da mesma família, «Quando é que um *libretto* é uma obra de arte literária?» ou «Quando é que uma partitura de um filme é uma obra de arte musical?». Exemplos de soluções positivas para estes dois últimos problemas são o *libretto* de Hofmannsthal para a *Ariadne auf Naxos* de Strauss e a partitura de Hans Werner Henze para o filme de Schlöndorff *Um Amor de Swann*.

Em primeiro lugar é preciso definir a sintaxe do termo «guião» de modo a saber se se pode falar de guião como um género e, no caso de uma solução negativa, a que género é que o guião vai ter que pertencer.

A solução negativa recomenda-se logo na base de uma distinção entre os conceitos de forma e género, de tal modo que faça sentido dizer, como no caso do soneto, que este pertence ao género lírico mas que não constitui *an sich* um género, mas antes um princípio de arranjo ou um princípio estrutural, usualmente pensado como forma. Nestes termos, a denotação de «género» não é a estrutura de uma obra de arte literária mas a totalidade formada por obras de arte literárias com diversas estruturas, mas com um predicado básico comum a todas elas.

É justamente a determinação deste predicado comum que leva à construção de uma árvore dos géneros, e como este predicado não é univocamente determinável a teoria literária tem assim, em momentos diferentes, sido levada a construir árvores diferentes.

Vale a pena contrastar a mais antiga, a que provém da *Poética* de Aristóteles, com uma árvore frequente na teoria literária contemporânea, para documentar essa oscilação na definição dos predicados básicos. Enquanto para Aristóteles a epopeia pertence ao género poesia, um género no qual Aristóteles apenas inclui também o drama ou a poesia dramática nas suas duas formas principais, a tragédia e a comédia, na teoria literária anglo-saxónica é frequente ver a epopeia como pertencendo ao género ficção, sendo os outros géneros desta árvore o drama e a poesia. Por poesia entende-se apenas poesia lírica, uma vez que a poesia dramática nesta classificação pertence ao género drama, enquanto à ficção pertencem não só a poesia épica, mas também o romance, a novela e o conto.

Não estou a afirmar que esta classificação anglo-saxónica dos géneros tem só vantagens, mas os seus inconvenientes são no entanto menores do que os de outras classificações propostas, como a de

Roman Jakobson, segundo a qual a definição dos predicados «lírico» e «épico» é feita apelando a uma postulada estrutura gramatical, a da primeira pessoa do singular no tempo presente e a da terceira pessoa do singular no tempo passado, respectivamente, uma estrutura que é, para mim, uma frívola ficção, uma vez que por ela ser-se-ia forçado a incluir no género épico todos os poemas líricos — são muitos — escritos na terceira pessoa do singular e no tempo passado. A este propósito seja-me permitido mencionar os adeptos do relativismo histórico, segundo os quais o conceito «género» é uma categoria literária do passado, a qual por volta do último *fin de siècle* deixou de fazer sentido. Mesmo concedendo a última parte do argumento, resta sempre a tarefa de definir os géneros do passado e regressam assim os problemas que se tinham tentado eliminar com o recurso ao relativismo histórico.

Um segundo momento, tão essencial como a definição dos predicados básicos é a definição da atitude a ter perante a classificação a propor. A teoria clássica do género deixa-se reconduzir a um único princípio, o da separação, o qual conduz na verdade toda a análise: na separação dos géneros quanto à pureza, na separação dos estilos, e na separação das classes sociais implicadas nos heróis da epopeia e da tragédia, até à mísera *plebs* da sátira e da comédia. É assim que este *pathos* da separação leva logo a distinguir a pureza da tragédia francesa da impureza da tragédia de Shakespeare e, acima de tudo, a exigir que a teoria do género seja de carácter prescritivo, não só por cada obra ter que pertencer a um género, mas também por esta pertença ser fixada num corpo de regras cuja execução se pode depois medir. O mesmo princípio da separação está presente na análise do estilo e na sua correlação com a origem social das personagens da obra de arte literária: o estilo alto ou elevado, que corresponde ao topo da escala social e é usado pelas personagens das tragédias e da epopeia, o estilo médio que corresponde à classe média e é característico das personagens da comédia, e o estilo baixo, associado como foi dito à sátira e à farsa.

Esta característica da separação hierárquica e regulativa da teoria clássica, presente em Aristóteles na inferiorização do poema épico perante o poema dramático, ultrapassou no século XIX as fronteiras da teoria literária e foi adoptado na estética musical sob a forma da separação da música absoluta da música de programa, em particular a concepção de que em si a sonata é uma forma mais pura do que o *Lied*.

Mas a experiência literária e musical nos dois últimos séculos não é conciliável com as posições da teoria clássica, e assim a concepção

segundo a qual os géneros enquanto universais existem *realiter* tem que ser abandonada e substituída por uma teoria que faça justiça à experiência. Em todo o caso, a recusa deste platonismo absoluto acerca da existência objectiva dos géneros como universais não deve conduzir à sua substituição por uma forma de nominalismo igualmente irrestrito, de acordo com o qual a literatura não é mais do que o agregado de poemas, de romances e de dramas que se convencionou descrever por meio de um nome comum a todos eles.

Entre o platonismo e o nominalismo a teoria filosófica que me parece capaz de produzir uma satisfatória teoria do género é a teoria dos jogos de linguagem. Segundo esta teoria não precisamos de conceber o género apenas como um nome ou um predicado e a existência objectiva do género pode ser redefinida em termos da existência objectiva das regras do jogo de linguagem que o género constitui. Definido o género como o jogo de linguagem torna-se também possível admitir a reformulação das regras de acordo com as necessidades ou novos objectivos dos utilizadores do jogo de linguagem, neste caso de «género». Nestas condições um género não é para ser comparado a um monumento do passado mas a uma instituição viva, capaz de reforma e de redefinição. É particularmente desejável que esta capacidade de reforma seja conciliada com a existência de um núcleo invariante, uma estrutura, de modo a ser possível ainda classificar as *Elegias de Duino* de Rilke como um poema lírico e os *Cantos* de Ezra Pound como poema épico.

Uma outra vantagem da teoria dos jogos de linguagem é proporcionar uma solução para a *vexata quaestio* «Como é que se garante que toda a obra de arte literária pertence a um género?», em virtude do facto de os jogos de linguagem não constituírem um conjunto fechado e ser assim sempre possível estabelecer um novo género. É instrutivo pensar que problemas derivados da falácia da intenção de Wimsatt, reformulada agora em termos da intenção de produzir uma obra para um género, são igualmente resolvidos pela circunstância de a teoria dos jogos de linguagem prescindir da intenção dos utilizadores do jogo de linguagem para que este se considere constituído e, nestes termos, não é teoricamente necessário nem documentar nem descobrir a intenção de género que o autor tinha ao escrever a obra para se definir o sentido dela.

Ficamos assim com a concepção de um género como uma estrutura objectiva mas não platonisticamente imóvel, convencional sem ser nominalisticamente arbitraria. Não só o número de todos os géneros

fica aberto como a atitude axiológica é substituída por uma atitude puramente descritiva. Não há géneros superiores e inferiores, assim como não há jogos de linguagem superiores e inferiores. O dogma da pureza do género também tem que ser abandonado, deste ponto de vista, uma vez que a confluência ou a intersecção dos jogos de linguagem é um facto da experiência, inconsistente com o princípio da separação. Resta o conjunto das regras, que fica comum às duas concepções, visto que um jogo de linguagem também tem que ser caracterizado por meio de regras.

Esta teoria é suficiente para estabelecer o guião como pertencendo ao género dramático, afastando assim a deformada concepção do guião como género. Mas uma vez esta pertença assegurada o guião partilha com todas as obras do género dramático do mesmo destino: ele só pode ser considerado uma obra de arte literária se a nossa percepção dele não depender da sua realização em filme, tal como *At the Hawk's Well* de Yeats é uma obra de arte literária mesmo que não seja realizada cenicamente. O facto de um realizador adoptar um guião para um filme não tem como consequência que o guião tenha deixado de existir como obra de arte literária, embora exista apenas como um fragmento da obra de arte filme. Estamos assim diante de uma situação em que o princípio de síntese da *Gesamtkunstwerk* é realizado com sucesso, mas ao contrário das exigências desta teoria temos agora uma síntese de duas artes autónomas, a qual é adequadamente perceptível na metáfora do contraponto a duas vozes, por vezes a cantar em uníssono, por vezes em movimento paralelo, por vezes em movimento contrário.

O cinema não produz só filmes de arte e não começou logo por ser uma arte, de modo que a teoria estética só se dirige ao fragmento da produção cinematográfica que cai debaixo do conceito de filme de arte. Estamos assim perante um complexo conjunto de relações que o filme de arte tem com as suas origens e, como vimos, com as outras artes.

4. De novo a metafísica

Para o senso comum as palavras «filosofia» e «metafísica» são vagamente sinónimas porque para este senso comum qualquer problema filosófico é em última análise um problema metafísico. A esta pretendida sinonímia acresce ainda o facto histórico de a metafísica ser, como a lógica, uma das disciplinas filosóficas já praticadas na Antigui-

dade Clássica e ter tido uma função estruturadora na teoria lógica do tratado de Aristóteles intitulado *Organon*. No seu perfil tradicional, a metafísica é a disciplina filosófica sublime *katexokén*. A teoria do ser e a reflexão sobre os conceitos de substância e existência são os pontos de partida da metafísica, por meio dos quais se enfrenta as questões fundamentais da perplexidade filosófica, como «O que há no mundo?», ou «Qual é a diferença entre o ser e o nada?». A estas questões em breve se juntaram as questões sobre a natureza da alma, quer a das suas relações com o corpo quer a da sua possível imortalidade. As grandes questões sobre a causalidade, a natureza dos conceitos de espaço e tempo, eram também do domínio da metafísica, a qual incluía igualmente o tratamento racional ou argumentativo do problema da existência de Deus, em particular uma avaliação das diversas tentativas de demonstração da sua existência.

Mas o desenvolvimento da ciência moderna veio suscitar um problema óbvio de divisão de competências e passou a ser gradualmente mais difícil defender a posição de que é a metafísica quem regula o que há a saber sobre a alma, o espaço, o tempo e a causalidade. Foi exactamente esta percepção que levou o filósofo francês do século XIX Auguste Comte a formular a sua Lei dos Três Estados, segundo a qual o conhecimento humano é um progresso por fases ou estádios, sendo o conhecimento obtido na metafísica superior ao obtido nas teogonias da Antiguidade mas ainda inferior ao conhecimento científico moderno.

Enquanto na tradição filosófica de língua alemã Kant foi o primeiro a relativizar o lugar sublime da metafísica na filosofia e no pensamento em geral, os seus principais adversários encontram-se no entanto entre os filósofos do Círculo de Viena. Sem querer entrar em pormenores de carácter histórico não posso deixar de mencionar dois textos clássicos que documentam a rejeição da metafísica pelo Círculo de Viena: em primeiro lugar o «Manifesto do Círculo»¹ e em segundo lugar o ensaio de Carnap «A Superação da Metafísica pela Análise Lógica da Linguagem».² No seu ensaio Carnap consegue provar que a teoria clássica da dedução é capaz de mostrar como problemas corren-

¹ Carnap, R., Hahn, H. e Neurath, O. *Wissenschaftliche Weltauffassung: Der Wiener Kreis*. Wolf, Viena, 1929.

² Carnap, R. «Überwindung der Metaphysik durch Logische Analyse der Sprache» in *Erkenntnis*, Vol. II, 1932. Trad. ing. de Arthur Pap, «The Elimination of Metaphysics Through Logical Analysis of Language» in Ayer, A. J. (org.). *Logical Positivism*. Greenwood Press, Westport, Connecticut, 1959, 1978.

tes da metafísica resultam do desconhecimento ingênuo da estrutura lógica da linguagem usada para o formular, de tal modo que uma análise desta linguagem é capaz de os revelar ou como falsos problemas ou como deduções falaciosas.

A dissolução física do Círculo de Viena e uma mudança de orientação quanto ao papel a desempenhar pela análise lógica levaram a que a metafísica recuperasse um pouco do seu prestígio mas não do seu lugar, uma vez que hoje se favorece — erroneamente ou não — uma disposição horizontal de todas as disciplinas filosóficas. O principal obstáculo à hegemonia da técnica da análise lógica na solução de problemas filosóficos foi sem dúvida criado por um interesse renovado — mais uma vez, erroneamente ou não — por aqueles aspectos da linguagem corrente que não podem ser explicados pela teoria lógica, tendo-se por isso passado a determinar o sentido de uma expressão ou de um conjunto de expressões por meio de um postulado «uso corrente», o qual é suposto agora substituir o sentido prescrito pela análise lógica. Nestas condições passa a ser possível reformular a metafísica, agora de um ponto de vista descritivo e não prescritivo. Qualquer sistema conceptual representado linguisticamente possui em estado latente um certo sistema de metafísica, um certo conjunto de ideias subconscientes acerca do ser, da alma e de Deus: a metafísica descritiva efectua, sem avaliar, uma representação desse sistema.

5. Negação, lógica e ontologia

Embora a inteligência humana se tenha desde muito cedo colocado diante da perplexidade que resulta do fenómeno puro da existência ou do ser, e se tenha também sem aparente sucesso posto à procura da resposta à pergunta «Porquê o ser e não o nada?», essa mesma inteligência nunca deixou de estar menos fascinada pela constatação da existência do raciocínio válido, ou da possibilidade de extrair uma consequência irrefutável a partir de uma proposição previamente dada.

Assim a inteligência humana é capaz de viver um fascínio duplo, ontológico como na pergunta pelo ser em vez do nada e lógico, como resulta do espanto pela existência da inferência válida, até no mundo mental da criança. Mas a inteligência é acima de tudo fascinada pela existência da inferência válida na vida mental do adulto, como se vê pela história da discussão em torno de paradoxos que surgem mesmo em níveis relativamente simples de conceptualização.

Um exemplo luminoso do espanto lógico é o que nos é dado pela consideração de uma actividade mental tão simples como a negação de uma proposição dada, porque neste caso um mínimo de análise é suficiente para mostrar que por baixo desta simplicidade está uma complexidade difícil de medir e que só no nosso século, de facto, foi formulada. Logo nos primórdios da lógica, na Grécia, se descobriu que a negação de uma verdade ou a afirmação da sua falsidade pode ter consequências perturbantes.

Se eu admitir que qualquer afirmação que faço só pode ser verdadeira ou falsa, o que parece ser razoável, vejo-me logo em dificuldades para determinar qual destes valores de verdade, verdadeiro ou falso, devo atribuir a certas proposições. A frase «Tudo o que eu digo é falso» é um bom exemplo desta dificuldade. Se eu disser que ela é verdadeira, então ela é falsa, visto que tudo o que eu digo é falso; mas se eu disser que é falsa, então ela é verdadeira, por ser isso justamente o que ela afirma e logo, o que eu estou a dizer agora, é verdadeiro.

Mas a negação é ainda uma fonte de perplexidades quando aparece associada ao já mencionado conceito de existência. Se eu afirmo «O Diabo não existe» e continuando a supor que esta minha afirmação ou é verdadeira ou é falsa, sou confrontado com diversas surpresas. Se suponho que é verdadeira, então «O Diabo existe» é falsa, o que parece também ser razoável; mas se suponho que «O Diabo não existe» é falsa, então «O Diabo existe» é verdadeira. Assim a falsidade de «O Diabo não existe» parece forçar-me a aceitar a existência do Diabo.

São estas perplexidades causadas pela negação e pela existência que estão na origem da constituição das disciplinas filosóficas mais antigas, a lógica e a metafísica, nas quais precisamente o espanto lógico e o espanto ontológico são articulados sob a forma de teorias capazes de satisfazer exigências básicas de racionalidade.

A lógica como disciplina autónoma começa com a distinção entre verdade e validade, em particular com a distinção entre a verdade de uma asserção ou de um conjunto de asserções e a validade do argumento ou do raciocínio em que elas ocorrem. É válido inferir da proposição «Alguns poetas são quadrúpedes» a conclusão «Alguns quadrúpedes são poetas» embora nem a primeira proposição, chamada premissa, seja verdadeira, nem a segunda, chamada conclusão, o seja. Mas não cabe à lógica determinar que proposições são verdadeiras ou não: à lógica cabe apenas dizer que um argumento é válido se de uma premissa verdadeira não se extrai uma conclusão falsa. Assim

a validade não é a mesma coisa do que a verdade, embora o conceito de verdade seja necessário para definir um argumento como válido.

A lógica mantém com a filosofia relações de um elevado grau de complexidade. Em primeiro lugar, como disciplina filosófica, a lógica é também uma experiência de surpresa ou de perplexidade, alguns traços das quais procurei esboçar com os meus exemplos. Mas a lógica separa-se em todo o caso da filosofia por se comportar de um modo diverso diante do teste dos resultados. Em contraste com a filosofia, a lógica constitui um corpo de resultados que não está sujeito a uma revisão constante, que tem um efeito cumulativo e que obriga toda a comunidade científica. Assim, as leis simples da inferência descobertas por Aristóteles ou as mais complexas descobertas por Frege representam uma forma de saber que não é típico na filosofia. Finalmente, põe-se a questão de saber se nestas condições é a lógica que deve determinar o curso de uma teoria filosófica ou se ao contrário é a filosofia que está na origem de uma teoria lógica.

M. S. Lourenço
Universidade de Lisboa

Um filósofo da evidência*

M. S. Lourenço

Universidade de Lisboa

Embora algumas posições filosóficas de Gödel sejam bem conhecidas, como o platonismo, a sua teoria do conhecimento é, em comparação, menos divulgada. A partir do «Problema da Evidência» de Hilbert-Bernays, I, pg. 20 seq., apresento a seguir os traços essenciais da posição de Gödel sobre a caracterização epistemológica da evidência finitista, com especial relevo para a história dos conceitos utilizados.

Embora Gödel tenha lido Filosofia desde o ensino secundário e depois na Universidade de Viena tenha frequentado o Círculo de Viena, o seu interesse profissional pela disciplina parece ter apenas acordado nos anos sessenta. Numa carta a Bernays datada de 1968, e numa referência ao seu trabalho de 1958¹, Gödel afirma que nesta data o seu objectivo predominante era de carácter matemático, em contraste com a posição em 1968, em que o conteúdo filosófico do problema da extensão veio ocupar o lugar do primitivo objectivo matemático.

Mas as posições filosóficas de Gödel assemelham-se aos enigmas daqueles filósofos da Antiguidade cujo pensamento só se conhece pela interposição de um segundo autor, o qual relata, com um inverificável grau de fiabilidade, as teses do primeiro. No caso de Gödel tem indubitavelmente esse papel Hao Wang, e é com o recurso aos seus livros que se pode reconstituir um esboço do que são as teses filosóficas de Gödel. O próprio Hao Wang tem uma judiciosa classificação dos escritos de Gödel em três áreas: uma área nuclear é constituída pelos resultados matemáticos de Gödel, e duas áreas adjacentes, uma com as *projeções* (para o futuro) de uma disciplina de Filosofia, na expressão de Brentano, como *ciência de rigor*, tipificada por uma teoria axiomática dos conceitos básicos da Metafísica, e uma outra de reflexões sobre temas heterogéneos como a origem da vida, a irrealidade

* Originalmente publicado no *Boletim da Sociedade Portuguesa de Matemática*, 55 (2006).

¹ Kurt Gödel (1958), «Über eine bisher noch nicht benützte Erweiterung des finiten Standpunktes», *Dialectica* 12, pp. 280-287.

do tempo segundo Leibniz e Einstein e o então novo círculo de problemas sobre a eventual mecanização do pensamento.

Para a identificação das teses de Gödel é assim preciso lançar mão não só dos seus escritos publicados e dos seus escritos não publicados, como também dos relatos dos seus alunos sobre a posição de Gödel em problemas filosóficos específicos, uma posição por vezes apenas articulada em entrevista ou em discussão com um restrito grupo de pessoas.²

É útil, antes de entrar propriamente na concepção de Gödel da caracterização do conceito de evidência, ter uma descrição de algumas das suas *outras* teses, em particular a definição de um método para a filosofia e o problema da união desta com a ciência.

Como segundo Gödel a filosofia actual está, para o ideal da ciência de rigor, no estado em que se encontrava a Física antes da obra de Newton, é de começar com a teoria (ou a *projecção*) da Análise Conceptual como o método da filosofia do futuro, da filosofia como ciência de rigor.

Uma primeira concepção da Filosofia como ciência de rigor e análise de conceitos apareceu na obra de Franz Brentano, cujo estilo filosófico era o oposto, quer pela forma reflectida quer pelo conteúdo preciso, do idealismo alemão. Para Brentano o uso da linguagem em filosofia não se deve afastar do uso da linguagem na ciência e ele foi, depois do ideal de Leibniz da *philosophia mathematica* e antes de Gödel, o mais enérgico defensor da tese segundo a qual o verdadeiro método da filosofia é o método da ciência e reciprocamente.

Extensionalmente um conceito é uma função (linguística) cujo domínio é um conjunto de objectos, chamado o domínio do conceito, e cujo valor é um único elemento do conjunto dos dois valores de verdade, verdadeiro e falso. *Intensionalmente* um conceito é o *sentido* do termo predicativo pelo qual o conceito é representado.

Os conceitos são o objecto em investigação na Análise Conceptual.

A função básica em Análise Conceptual é para Gödel a *percepção de conceitos*, os quais são susceptíveis de ser conhecidos com o mesmo grau de fiabilidade com que conhecemos os objectos do mundo exte-

² Gödel teve consciência de estar em oposição ao «espírito do tempo» (ao que ele chama *Zeitgeist*) e recusou por isso dar a conhecer as suas teses em Filosofia, por recear ter um público desfavorável. Mas com Bernays, que conhecia com detalhe a filosofia de Brentano e de Husserl, Gödel teve discussões prolongadas ainda antes de 1958. (Brentano exerceu a sua influência também em Zurique, onde teve uma cátedra desde a sua emigração de Viena até 1918).

rior. A percepção de conceitos pode ser ensinada e é possível corrigir gradualmente o grau de fiabilidade com que um conceito é percebido. O instrumento adequado para essa correção é justamente a *história conceptual*, onde a *evolução* do conceito é investigável. Assim os conceitos descobertos pela nossa percepção evoluem à medida que a própria percepção evolui.

Exemplo: A proposição «Todos os conjuntos são construtíveis» foi numa ocasião considerada por Gödel como um axioma e mais tarde considerado como sendo uma proposição falsa.

A finalidade da Análise Conceptual é a transformação de um conceito vago num conceito preciso. Há diversos paradigmas da Análise de Conceitos e embora Gödel tenha favorecido sempre como paradigmas a análise de Dedekind do conceito de «número» e a análise de Turing do conceito de «função calculável», é óbvio que Gödel 1958 pode também servir como paradigma de análise, neste caso do conceito de função calculável de tipo finito.

Uma Análise de Conceitos é representada num sistema de axiomas: se o sistema de axiomas tem um modelo então os conceitos que ocorrem no sistema são precisos; *a fortiori* se uma teoria axiomática é categórica, os conceitos analisados nos axiomas são conceitos precisos.

Existe uma diferença gradual entre a percepção de pequenos conjuntos de números e a percepção dos números naturais como uma totalidade acabada. A *idealização* é o instrumento necessário para passar de uma para a outra forma de percepção. Gödel concebe a idealização como sendo uma função de síntese, tal como esta é definida por Kant na *Crítica da Razão Pura*, onde o conceito de uma totalidade (infinita) é identificado com a representação da síntese perfeita das partes, uma vez que esta síntese não pode vir da percepção do todo infinito (uma percepção que por definição não existe), com a reiteração desta síntese das partes até ao infinito. A percepção do conjunto dos números naturais como uma totalidade é assim o resultado do exercício da actividade de síntese, melhor definida no § 18 da *Crítica da Razão Pura* como sendo *a unidade transcendental da apercepção*, a qual unifica no conceito de *um objecto* tudo o que na intuição nos é dado como uma multiplicidade.

A estabilidade da teoria clássica dos números garante a objectividade das nossas percepções relativamente aos números naturais. É claro que o conceito de percepção tem que ser generalizado de modo a incluir também a *percepção de conceitos* e em particular a percepção de conceitos a partir da *perspectiva certa*, como foi exemplificado pela

definição de Turing do conceito de «processo mecânico». É este género de percepção que Gödel designa por *intuição matemática*.

Mas a intuição matemática de Gödel é diferente da evidência intuitiva de Hilbert. Hilbert tinha-se apoiado na definição bipartida de Kant na *Crítica da Razão Pura* (113, 175) em particular na sua análise da intuição sensível.³ Esta é apresentada em duas formas, a pura e a empírica e enquanto ambas se desenvolvem no quadro espaço-tempo, só a intuição empírica se relaciona com um objecto por meio de uma sensação.

Nestes termos, a percepção de objectos de segunda ordem, como conceitos ou como os funcionais recursivos de Gödel 1958 não caem dentro dos limites da evidência intuitiva de origem kantiana. Em última análise as definições de Gödel são incompatíveis com a adopção das teorias causais gémeas, a teoria causal da percepção e a teoria causal do conhecimento, uma vez que não se pode supor que objectos de segunda ordem actuem causalmente sobre o sujeito cognitivo.⁴

Os conjuntos e os conceitos são as entidades básicas da ontologia de Gödel. *Conjuntos puros* são os que se obtêm começando com o conjunto vazio e depois reiterando a operação «o conjunto de...».

Os conceitos são funções e relacionam-se com os conjuntos através da extensão do conceito. A relação de «pertença» é para um conjunto o que a relação de «estar na extensão de...» é para um conceito. Os conjuntos e os conceitos são ambos aspectos da realidade objectiva. Os números são também parte da realidade e assim num termo como «as 7 cores do arco-íris» fazem parte da realidade objectiva não só as cores (vermelho, amarelo, etc.) como também o número 7.

À existência objectiva dos objectos matemáticos corresponde a objectividade do nosso conhecimento acerca deles. Esta objectividade exprime-se num conceito de verdade para o qual o *Tertium non Datur* é válido ou, equivalentemente, na convicção de que, por exemplo, a Conjectura de Goldbach é verdadeira ou falsa. Assim os objectos

³ A tradução do termo de Kant «Anschauung» por «intuição» é desvantajosa, porque na linguagem corrente o sentido do termo «intuição» é em geral compreendido como «ter sorte a adivinhar».

⁴ Para uma tentativa de conciliação ver P. Maddy, «Perception and Mathematical Intuition», *The Philosophical Review*, LXXXIX, N.º 2, (April 1980). Deve-se no entanto ter em conta que o projecto de P. Maddy depende da concepção de conjunto como um «indivíduo» e não como um objecto abstracto.

matemáticos não são apenas nomes ou construções arbitrárias do sujeito cognitivo.

Mas embora nós tenhamos uma percepção destes objectos, as nossas ideias sobre as suas propriedades são por sua vez formadas na base ainda de algo muito diferente, nomeadamente na base de um *conjunto de verdades* que nos é imediatamente dado, o qual Gödel qualifica com o termo tradicional de «evidente».

Na literatura filosófica o termo «evidência» é usado em dois sentidos diferentes: num primeiro sentido, corrente em filosofia da ciência, o termo cobre o conjunto dos objectos e das relações que apoiam uma hipótese científica e ocorre neste sentido em expressões como «evidência empírica» ou «evidência estatística»; num segundo sentido, o sentido usado por Hilbert-Bernays e por Gödel 1958, o termo provém da tradição da teoria escolástica da intencionalidade e da existência de cognição intuitiva, uma teoria desenvolvida essencialmente pelos lógicos de Oxford no século XIV, o platonista Duns Escoto e o nominalista William Ockham, e a qual é proposta para explicar a experiência de uma verdade como indubitável, ou como uma evidência. Mas a análise da evidência de Hilbert-Bernays, que Gödel retoma no seu ensaio 1958, é um estágio posterior da teoria neo-escolástica de Brentano, o qual adapta as posições de Escoto na intencionalidade, e de Escoto e Ockham na evidência, a uma ontologia anti-realista.

Para identificar a intencionalidade⁵ dos actos da mente Duns Escoto parte do facto de que somos capazes de reflectir sobre os nossos próprios actos e que obtemos assim, através desta auto-reflexão, a concepção de um sujeito cognoscente. Este acto de auto-reflexão é o resultado de uma função da mente. Assim é possível conhecer os objectos *visados* pelos actos da mente, isto é, a sua intencionalidade, mesmo quando estes objectos estão no momento ausentes da mente. Diz-se então que estes objectos são *mentalmente inexistentes*, mas alcan-

⁵ «Intencionalidade» denotará a seguir uma relação binária entre um estado mental e o seu objecto (ou conteúdo). Exemplos: (O estado mental em itálico) 1) O que Frege *queria dizer* com «a estrela da manhã» era a deusa do amor e não o planeta Vénus; 2) Uma pessoa *A quer dizer* com «virtude» o que uma pessoa *B quer dizer* com «vício». As suas propriedades essenciais são: i) a relação é dada pelo sentido fregiano dos termos, i.e., pelo seu modo de apresentação; ii) a relação obtém sem implicação existencial para o 2.º argumento; iii) a expressão da relação cria um *contexto intensional*, logo sem substituição *salva veritate*.

çáveis intencionalmente. Obviamente este conhecimento só é possível se se postular a existência do acto de auto-reflexão. Na verdade a existência deste acto é-nos dada na experiência e o *facto* da auto-reflexão implica a existência da função mental que a realiza e a que se chama *intuição*. Assim o acto de auto-reflexão ensina-nos que possuímos conhecimento de propriedades de objectos as quais são reconhecíveis por uma função da mente, a intuição, mesmo quando estes objectos são mentalmente inexistentes.

Para Duns Escoto a experiência da intuição mostra-nos que podemos alcançar o conhecimento de objectos abstractos (ou universais). Na verdade podemos conhecer um conceito e alguns dos seus aspectos, como o seu *sentido*, de um modo completamente diferente do que conhecemos os objectos das nossas percepções sensíveis e sem qualquer intervenção de uma faculdade dos sentidos. A experiência intuitiva também nos revela a existência de relações que obtêm entre objectos, mesmo quando estes objectos não são percebidos pelos sentidos. Temos sobretudo um conhecimento intuitivo do próprio acto por meio do qual conhecemos estes diferentes objectos, *como um acto interno*, que tem lugar na nossa própria mente. E assim vemos a dar o nosso assentimento a juízos imunes a dúvida ou a contra-exemplo, quando por exemplo aceitamos axiomas como evidentes.

Foi Duns Escoto quem primeiro fez a distinção entre *cognição intuitiva* e *cognição abstractiva*: com a primeira predica-se o conhecimento acerca de um objecto *existente*, (por exemplo, com coordenadas espaço-tempo), pela segunda a existência do objecto não é parte do conhecimento predicado pelo termo. A cognição intuitiva pode predicar o conhecimento não só acerca dos objectos (materiais) do mundo exterior como também acerca dos objectos (imateriais) da actividade mental. Mas a cognição abstractiva não pode garantir a certeza indubitável dos nossos juízos sobre objectos existentes, quer os materiais no mundo exterior quer os imateriais no mundo interior. Na verdade só a cognição intuitiva torna possível a introspecção como intuição dos objectos do mundo interior e a evidência como um predicado de um juízo verdadeiro.

Para um nominalista como Ockham, de acordo com o princípio de que o todo é apenas a união das partes, o conhecimento de uma proposição verdadeira é considerado como evidente se e somente se o modo de conhecer o sentido dos termos produz o modo de conhecer o sentido da proposição (verdadeira). Assim é *o modo* (não o conteú-

do) como os termos de uma proposição são dados a conhecer que determina se a proposição é ou não evidente.

A posição de Ockham ilustra-se bem através da distinção entre proposições necessárias e contingentes. O resultado geral é que uma proposição necessária é evidente mas que nem todas as proposições evidentes são necessárias. Se uma proposição é necessária qualquer das formas de conhecimento (intuitivo ou abstractivo) é suficiente para garantir a evidência da proposição. Mas no caso de uma proposição contingente só o conhecimento intuitivo dos termos da proposição produz a evidência da proposição. Assim se se abstrair da questão da existência ou não existência do objecto (de conhecimento), segue-se que não se pode ter conhecimento evidente de asserções existenciais acerca deste objecto. Logo para explicar o facto de que consideramos algumas proposições existenciais como evidentes é necessário admitir a existência de uma outra forma de conhecimento, o conhecimento intuitivo. O conhecimento intuitivo é uma relação directa entre a mente e o objecto, sem recurso à mediação por um conceito. É para Ockham geneticamente a primeira forma de conhecimento. Como para um nominalista as únicas coisas reais são indivíduos, o conhecimento começa com o conhecimento de indivíduos. Não se parte do universal para o indivíduo; começa-se com indivíduos e determina-se a seguir em que sentido é que os conceitos que obtemos com eles são universais. Uma forma diferente de conhecimento é obtida pela cognição abstractiva, a qual permite passar do conhecimento de indivíduos para o conhecimento de um conceito (ou um universal) pela exclusão de juízos de existência (ou de diferenças individuais). Mas tal como Duns Escoto, Ockham pensa que a cognição intuitiva produz um grau mais elevado de evidência.

No século XIX a polémica contra a caracterização de Kant do juízo aritmético como um juízo sintético foi bem representada por Franz Brentano, o qual concebeu e realizou o projecto de uma definição, a que ele chamou empirista, do conceito de verdade.

Como para Brentano a introspecção é um método fiável de investigação, pode-se generalizar o conceito de *experiência* de modo a incluir os resultados da introspecção, sob a forma de uma descrição dos fenómenos da mente,⁶ em particular dos actos mentais. A relação entre um acto da mente e o seu objecto faz daquele um *acto intencio-*

⁶ Equivale por isso ao sentido do termo inglês *experience*.

nal. Actos mentais em particular e fenómenos mentais em geral são para Brentano definidos como aqueles fenómenos que *referem* de um modo intencional (ou apenas mental) um conteúdo ou um objecto, mesmo quando está mentalmente inexistente.

Um acto mental distingue-se dum acto físico, (como pregar um prego na parede), pelo facto de a relação entre o acto físico e o seu conteúdo ser uma *relação externa* e a relação entre um acto mental (como duvidar que Pedro negou Cristo) e o seu conteúdo ser uma *relação interna*. Em particular a intencionalidade de um acto mental é constituinte do acto, a sua alteração produz um outro acto.

Exemplo: «Duvidar que Pedro negou Cristo». Negar Cristo está com Pedro numa relação interna, porque negar Cristo é constituinte da identidade de Pedro.

Isto é na verdade um corolário da doutrina de Leibniz de que um objecto é determinado pela soma dos seus predicados. Assim para ver que negar Cristo está com Pedro numa relação interna basta provar que não podem existir *duas* situações com *os mesmos* objectos, a situação S_1 , em que Pedro nega Cristo e a situação S_2 em que Pedro não nega Cristo. (No exemplo o Pedro de S_2 é diferente do Pedro de S_1 , visto ser igual a «o Pedro de S_1 menos negar Cristo» e o Cristo de S_2 é diferente do Cristo de S_1 , visto ser igual a «o Cristo de S_1 menos a negação de Pedro».)

Em geral defende-se que uma tal relação é uma relação necessária, expressa pela implicação «se Pedro nega Cristo então necessariamente Pedro nega Cristo».⁷

Os actos mentais são descritos por verbos transitivos como por exemplo «pensar que P », «constatar que P », «duvidar que P », etc. e *aquilo* que se pensa, se constata ou se duvida é o conteúdo intencional do acto de pensar, constatar ou duvidar. Assim para Brentano constatar a evidência da verdade de P é equivalente a aceitar que P é irrefutável por meio de um contra-exemplo.

Brentano propõe uma classificação dos juízos, rival da classificação de Kant, de modo a separar três classes fundamentais: juízos formados a partir de uma percepção, juízos formados a partir de uma memória e os juízos incorporados em axiomas.

⁷ Contra relações internas em geral e a sua interpretação modal em particular ver G. E. Moore, «External and Internal Relations», *Proceedings of the Aristotelian Society*, 1919-1920.

Como há duas formas de percepção para Brentano, a percepção dos fenómenos físicos e a percepção dos fenómenos da mente, os juízos formados a partir de percepções dão origem a duas subclasses correspondentes. Mas só a verdade dos juízos formados a partir da *percepção* dos fenómenos da mente e a verdade dos axiomas pode ser considerada como evidente. A verdade de um juízo formado a partir de uma *memória* não pode ser considerada como evidente.

Adaptando a distinção de Leibniz entre verdades de razão e verdades de facto, Brentano distingue a *evidência apodíctica*, que resulta dos axiomas através da análise do sentido dos conceitos que ocorrem no juízo verdadeiro, da *evidência imediata*, como a que é obtida na introspecção. Entre as duas formas de evidência existe a seguinte relação: a reflexão sobre os dados da evidência imediata transforma esta em evidência apodíctica.

Exemplo: a proposição «A representação (na mente)⁸ de um par de objectos é condição necessária de um juízo sobre eles» é tipicamente um juízo apodíctico mas os termos «representação» e «juízo» são para Brentano dados da introspecção.

Husserl, discípulo de Brentano, foi o primeiro crítico da doutrina da evidência apodíctica como irrefutabilidade por meio de um contra-exemplo. Para Husserl o facto de um contra-exemplo não ser conhecido é simplesmente contingente, pelo que o carácter apodíctico de um juízo verdadeiro tem que ser derivado da *percepção da estrutura* que os conceitos que ocorrem no juízo representam. A evidência apodíctica de uma proposição tem que resultar da análise do papel dos conceitos para a definição da estrutura. E é esta análise (e não a inexistência do contra-exemplo) que produz a evidência apodíctica com que a verdade da proposição é conhecida.

Deve-se reparar que um conceito como «percepção da estrutura» pertence ao vocabulário do platonismo de Husserl e é por isso inconsistente com o programa empirista de Brentano, de recondução do conhecimento à experiência.⁹

⁸ «Representação» pretende ser uma tradução do termo de Brentano *Vorstellung*. Há uma disparidade de traduções deste termo, importante no vocabulário de Brentano e no vocabulário finitista de Hilbert. Os objectos em sentido finitista têm que ter *Vorstellbarkeit*, pelo menos em princípio, i. e., têm que ser representáveis (na mente).

⁹ A posição de Gödel em «O que é o problema do Contínuo de Cantor?» é obviamente congruente com a de Husserl. Os axiomas da teoria dos conjuntos, diz Gödel, impõe-se-nos como verdadeiros, constituem por isso um exemplo de

Para a extracção do significado filosófico de uma teoria matemática dispõe-se essencialmente de duas linguagens: a linguagem da ontologia, com a qual se analisa a questão da existência e a da natureza dos objectos da teoria, e a linguagem da epistemologia, com a qual se determina o género e o grau de conhecimento obtido na teoria.

Gödel, em textos como «A Lógica Matemática de Russell» e «O que é o problema do Contínuo de Cantor?», situa-se num ponto de vista ontológico, dentro do qual toma posição a favor do platonismo ou da existência real dos objectos matemáticos. Mas mesmo quando se adopta um ponto de vista unilateralmente ontológico como forma preferida de análise, problemas sobre evidência colocam-se inevitavelmente, quer durante a descoberta de novos axiomas, quer na justificação de axiomas existentes, quer na reflexão e na arbitragem de axiomas duvidosos. Em particular, um conceito adequado de evidência deve permitir responder ao seguinte catálogo de problemas:¹⁰

- P1 Numa demonstração existem diferenças interessantes entre os símbolos no papel e os pensamentos que eles representam?
- P2 Se se foi convencido por uma demonstração indutiva de uma proposição $(\forall n) 2^n + 2^n = 2^{n+1}$, é a evidência a que se faz apelo na demonstração finitista?
- P3 Em que é que se pensa quando se pensa a^n ?
- P4 Compreender uma demonstração como uma demonstração Δ (Δ pode ser substituído por «intuicionista», «formalista», ou «finitista»). É empiricamente decidível se se está consciente de compreender uma demonstração como uma demonstração Δ ?
- P5 Existe um *acto de compreensão* Δ ?

O ponto de vista da epistemologia¹¹ é tomado por Gödel essencialmente nos seus trabalhos de 1933,¹² sobre a aritmética intuicionista,

evidência apodíctica. Este facto permite inferir a existência subjacente de uma percepção, uma vez que se não tivéssemos uma percepção dos objectos da teoria dos conjuntos não faria sentido constatar que, para tais objectos, os axiomas se impõem como verdadeiros. Este conceito de uma percepção inferida é uma construção por analogia, a partir do conceito de percepção de um objecto material.

¹⁰ Soluções para P1 e P3 em Kreisel 1967.

¹¹ O ponto de vista epistemológico (e o seu vocabulário básico) foi introduzido em Hilbert-Bernays, I, 1934, onde é descrito como «O Problema da Evidência» (*Das Evidenzproblem*) o qual consiste em avaliar a diferença de evidência que resulta quando se transita de percepções concretas para reflexões abstractas. Para satisfazer

1942 sobre os *limites* do conceito intuicionista de evidência, 1958-1972 sobre uma extensão do ponto de vista finitista por meio de funcionais recursivos. É nestes trabalhos que é posta em acção a teoria (implícita) da evidência de Gödel. Embora não seja possível explicitar completamente uma tal teoria, é possível no entanto reconhecer que a concepção de Gödel se orienta pelos princípios seguintes:

- i) A meta a alcançar por uma teoria matemática é para além da verdade a evidência.
- ii) Existe uma hierarquia de teorias por graus de evidência.¹³
- iii) A evidência de uma teoria diminui à medida que se introduz quantificação universal sobre classes de objectos intensionais (exemplo: «Todas as demonstrações intuicionistas tais que $F(d)$ ».)

Os trabalhos de 1958 e 1972 têm o mesmo objectivo estratégico, *em nome de um acréscimo de evidência* traduzir a aritmética intuicionista no sistema T , mas diferem um pouco na execução. Em 1972 os termos especialmente interessantes de 1958 para a filosofia da evidência foram eliminados¹⁴ e, embora o novo texto não altere a concepção

o ponto de vista finitista em termos de evidência tem que se exigir que numa definição finitista, i) os objectos sejam sempre representáveis (na mente) e que ii) os processos sejam sempre realizáveis. Kronecker satisfaz o ponto de vista finitista mas o Intuicionismo já é uma extensão.

¹² «Zur intuitionistischen Arithmetik und Zahlentheorie», *Ergebnisse eines mathematischen Kolloquiums*, 4, 1933.

¹³ Só dentro da teoria finitista se podem reconhecer os seguintes (por ordem descendente): A) Uma construção que não depende de condições e que se pode visualizar imediatamente como uma totalidade acabada. Exemplo: $2 + 2 = 4$. B) $10^3 + 2 = 1002$. Aqui já não se exige a construção real de 10^3 ; bastará mostrar que $10^3 + 2 = (1000 + 1) + 1 = 1002$. C) $a + 2 = 2 + a$. Já é uma proposição universal e assim interpretada como sendo a asserção: «é conhecido um método geral de construção o qual aplicado a um número natural já construído a permite a construção da igualdade». D) Provar propriedades de números naturais a partir de uma condição da qual não se sabe se pode vir a ser satisfeita.

¹⁴ Embora Gödel declare na nota *a* de 1972 que o texto inglês não é uma tradução literal, algumas das liberdades de tradução constituem um problema interpretativo: i) pp. 272, 273, 274, «concrete intuition» ocorre como tradução da expressão hilbertiana *anschauliche Evidenz*. Ainda que se possa aceitar «concerte» como tradução de *anschaulich*, não se pode aceitar «intuition» como tradução de *Evidenz*, uma vez que a intuição é uma faculdade mental e a Evidência é uma propriedade de um juízo verdadeiro. ii) p. 275, a definição crucial foi reformulada tendo sido eliminada

original, nesta o acto de compreensão necessário para compreender a definição de funcional deixa de ser explícito. Mas esses termos são essenciais para decidir se é uma questão empírica determinar se se está consciente de estar a compreender uma definição *como* uma definição construtiva (Problema 4, acima). É claro que uma resposta positiva implica a aceitação da existência do *acto de compreensão* construtivo (Problema 5).

Em 1958 Gödel confronta-se com questões de evidência em dois momentos, um primeiro momento durante o relato e a análise do *insight* de Hilbert-Bernays, segundo o qual a evidência intuitiva é insuficiente para a demonstração da consistência da aritmética e que tem por isso de ser suplementada pela evidência abstracta e, num segundo momento, durante o esboço da arquitectura do sistema *T*, com a separação entre o acto intencional da compreensão do sentido do conceito de função calculável de tipo finito e a evidência dos axiomas propostos para *T*.

O significado histórico de 1958 reside no facto de Gödel reabilitar o conceito tradicional de evidência, mesmo perante um ambiente circundante hostil.¹⁵ Em 1958 a epistemologia e a filosofia da mente então dominantes eram em primeiro lugar, desde os *Principia Mathematica*, contra a existência de relações internas, em segundo lugar, desde o Círculo de Viena, a favor da eliminação de relações intensionais na teoria unificada da ciência e, em terceiro lugar, desde as *Investigações Filosóficas* de Wittgenstein a favor de um behaviorismo encapotado.

Mas logo a definição (informal) de evidência abstracta de Bernays, adoptada por Gödel em 1958, subscreve a adopção de um princípio de intensionalidade, implicando assim o conhecimento de objectos mentais, como demonstrações intuicionistas ou o sentido de proposições, bem como a existência de relações internas e a fiabilidade da intuição.

Assim o seu trabalho de 1958 não se insere na epistemologia oficial do Círculo de Viena, onde Schlick e Carnap consideraram o concei-

a referência ao acto de compreensão dos conceitos «função computável de tipo t_i » ($i = 0, 1, \dots, k$). A frase da linha 6 da mesma página «this general fact is constructively evident» é ambígua: o que tem que ser construtivamente reconhecível é que a operação é sempre possível. iii) p. 276, os axiomas deixam de ser apresentados como evidentes.

¹⁵ A título de comparação: em 1994, *The Oxford Dictionary of Philosophy*, de Simon Blackburn, não tem uma entrada «evidência», apesar da abundante escolha de temas em epistemologia.

to de evidência como destituído de sentido e uma faculdade de intuição como inexistente, como também não se insere nem na epistemologia naturalista de Quine nem na filosofia da mente de Wittgenstein. A linha de investigação epistemológica que através de Bernays chegou a 1958 provém da orientação anti-kantiana e neo-escolástica dominante no Império Austríaco associada com os nomes de Franz Brentano e mais tarde de Edmund Husserl.

M. S. Lourenço
Universidade de Lisboa

Bibliografia

- Bernays, P. 1935. Sur le Platonisme dans les Mathématiques. *L'Enseignement Mathématique*, 34.
- 1959-1960. Ludwig Wittgensteins 'Bemerkungen über die Grundlagen der Mathematik.' *Ratio* 2.
- Brentano, F. 1874. *Psychologie vom empirischen Standpunkt*, Leipzig.
- 1930. *Wahrheit und Evidenz*, Leipzig.
- 1929. *Über die Zukunft der Philosophie*, Leipzig.
- Brouwer, L. 1948. Consciousness, Philosophy and Mathematics. *Proc. Xth International Congress of Philosophy*, Amsterdam.
- Chisholm, R. 1960. *Realism and the background of phenomenology*. Illinois.
- Day, S. 1947. *Intuitive Cognition, A Key to the Significance of the later Scholastics*. New York.
- Dummett, M. 1993. *Origins of Analytical Philosophy*. Harvard University Press.
- Gilson, E. 1952. *Jean Duns Scot*. Paris.
- Gödel, K. 2001. *Collected Works*, ed. S. Feferman et al. Oxford.
- Hilbert-Bernays. 1939. *Grundlagen der Mathematik*, vol. I. Berlin.
- Husserl, E. 1891. *Philosophie der Arithmetik*, Halle.
- 1938/1939. Die Frage nach dem Ursprung der Geometrie als intentionalhistorisches Problem. *Revue Internationale de Philosophie*, vol. 1.
- Kneale, W. & M. 1972. *O Desenvolvimento da Lógica*. Trad. M. S. Lourenço, Lisboa.
- Kreisel, G. 1967. Mathematical Logic: what has it done for the philosophy of mathematics? *Bertrand Russell, Philosopher of the Century*, ed. R. Schoenman, London.
- 1970. Die formalistisch-positivistische Doktrin der mathematischen Präzision im Lichte der Erfahrung. *Zentralblatt für Mathematik und ihre Grenzgebiete*, 196.
- Ockham, W. 1675. *Summa Totius Logicae*, Oxford. Ed. moderna por P. Boehner, New York, 1945.
- Wang, H, 1988. *Reflections on Kurt Gödel*. Massachusetts Institute of Technology.

The axioms of wisdom

M. S. Lourenço
University of Lisbon

Following a suggestion of Brouwer an axiomatic definition of wisdom is proposed.

I. In order to isolate the axioms of wisdom we have to separate wisdom from religion and from philosophy.

This follows from the fact that wisdom provides a measure by means of which we can classify a religious doctrine as expressing wisdom, respectively unwisdom. Today it is almost a tautology to say that not every religious doctrine expresses wisdom.

On the other hand, there are several examples of philosophical positions that do not express wisdom, in spite of the fact that the Greek word 'philosophy' implies the love of wisdom. It is again a tautology to say that today not every philosophical doctrine as such expresses wisdom.

II. To perceive the axioms of wisdom we first have to assume *the universal constant of meaning* as a primitive concept.

Thus the axioms of wisdom are:

1. **Intentionality:** A theory expresses wisdom if and only if the underlying concept of meaning is intentionality.
2. **Holism:** A theory expresses wisdom if and only if in the underlying theory of meaning an element has no meaning in isolation but depends on the meaning of the totality itself.
3. **Grace:** A theory expresses wisdom if and only if its truth is a discovery that does not depend on the properties of the symbolism used.
4. **Actual Infinity:** A theory expresses wisdom if and only if its truth implies the actual existence of a non-denumerable infinite totality.

M. S. Lourenço
University of Lisbon

What is structuralism?*

M. S. Lourenço
University of Lisbon

I. Introduction

Good afternoon, ladies and gentlemen. Let me start by saying how I interpret this meeting, because that does determine what I am going to say next. I think about this meeting as a kind of technology fair, where products are for the first time shown and tested. The purpose of this meeting, as I see it, is to bring together professional philosophers doing analytic philosophy and ask them to show their products, so that hopefully we end up not only with a map of who is doing what but also with an increased pride in having given an old profession a distinctly new profile.

I would like to say why I think this is important and for that purpose I will revert to my technology metaphor. The practice of analytic philosophy requires such an intensive training and such a peculiar skill that each philosopher becomes by its practice a specialist, in Taylor's sense of the word. This is a reversal of a long-standing tradition of the philosopher as a generalist and, as we all know, anybody who wants to do analytic philosophy has to give up any claims to being a generalist. However, this is not to be seen as unfortunate in any way, because it is simply a by-product, on one hand, of the size of the literature in each of the traditional philosophical domains and, on the other, of the modern diversity of methods used in each of them. But we all eventually reach a point where we have to recognize that there is something unprofessional about mastering everything.

* This text was presented by M.S. Lourenço as the inaugural lecture of the First Meeting of the Portuguese Society for Analytic Philosophy, ENFA1, in Coimbra 2002. It was decided to include this lecture in the present volume even though the author did not have the opportunity to revise it for publication, and that explains some imprecisions in the text.

On the other hand, I hope to be able to use the tools of analysis to rebalance the alleged disadvantage of a relative loss of scope. This analysis will show that it is, at least in principle, always possible to reach an overview, because philosophical theories split up in what I will call local theories, i.e., domain variant and global theories, i.e., domain invariant. A conspicuous instance of this dual structure is the use of global theories, like logic, epistemology and ontology, in individual domains, so that it becomes natural to define a theory in Ethics as ethical realism or ethical idealism and a faulty reasoning in Aesthetics is called the intentional fallacy.

And so our work has to proceed towards a reconciliation of the specialist's intense focus with the attainment of a balanced overview.

II. Open Questions and Strategic Options

In that sense, I propose, as my contribution to this Meeting, to show the line of goods I am interested in. But in my lecture I shall not be really engaged in problem solving, because I will rather attempt to answer the following heuristic question: what makes it exciting for a philosopher to engage in structuralist philosophy?

As I share Paul Benacerraff's view that philosophy of mathematics is philosophy in its most abstract state, I will try to consider general reasons that a philosopher might be inclined to use and to accept as grounds for an engagement in structuralism.

I would like to identify two sets of reasons a philosopher might accept as grounds for such an engagement:

- (i) The intrinsic interest of the open questions and
- (ii) The obvious thrill in defining parametric or strategic options.

I start with the first set. I provide a short list of the best-known open questions so that I may say something as to why I find them interesting.

Question # 1: The varieties of structuralism. Not only does Structuralism divide along the lines between categorical and modal logic, but there are also several structuralist positions on themes like truth, reference and ontology. What are we to make of this diversity? The structuralist answer to this question is that this diversity is no more damaging to an understanding of the structuralist position than

the diversity of definitions of usual objects, like natural numbers, functions and sets, is to our understanding of these objects.

Question # 2: The transformations of the Identity Problem.

This problem had an initial treatment by Benacerraff in 'What Numbers Could Not Be,' which articulates the classical structuralist position, and it then migrated to the theory of patterns in the form of the disputed claim that patterns are not in the field of the identity relation. Should there be a definition of identity for patterns? If yes, then in what form?

Question # 3: The relation between the concepts of categorical set and pattern.

Contrary to what was usual in the history of science, here the mathematical formulation preceded the philosophical. This fact is not only a counterexample disproving old time Positivism (Comte's law) but it also forces an answer to the McLarty question as to whether structuralism is fulfilled or obviated by the mathematical theory, in the sense that there is no need for a philosophical theory of abstract structures.

Question # 4: In what sense is the concept of a category a foundational concept?

An answer to this question depends on what one is supposed to recognize as a foundation and in this direction there is already J.P. Marquis's perceptive analysis of the relation 'S is a foundation for T,' where S and T range inter alia over mathematical theories. Since he is not satisfied with his own solution, we have good reason to look for a way to complement it.

I call a strategic option also a parametric option because I intend to refer with this term to a goal or a value that remains unchanged or unquestioned throughout the development of a theory. Example: formal evidence will be chosen as a goal, because it is assumed to be more reliable than intuitive or informal evidence.

In this sense, what are then the most pressing strategic options in Structuralism?

Strategic Option #1: Abstractionism. A commitment to abstractionism entails a redefinition of the doctrine and the rejection of Dummett's formulation of abstractionist concept formation.

Here is Dummett's description of abstractionist concept formation. His main thesis is that if abstractionism were to be coherent

there would be a unique formulation of abstractionist concept formation. But in fact for him there are three different abstractionist theories and therefore abstractionism is incoherent. Dummett orders the theories as follows:

Theory #1: It is possible to attain an as yet new concept by contemplating a number of otherwise distinct objects.

Theory #2: The attainment of the new concept (in #1) is effected by abstracting from the properties differentiating the objects in question. This is achieved by diverting attention from those properties.

Theory #3: The operation of abstraction (in #2) can also generate abstract mental constructions, that is, abstract objects or structural objects. These lack all those properties abstracted from and have no others in their place.

We cannot accept that this formulation does any kind of justice to the abstractionist position, let alone that it shows that there are three different abstractionist theories.

If one analyses Theory #1 one soon finds out that it actually has no content, since there is no satisfactory way to define the concept of ‘contemplation of otherwise distinct objects.’ Turning to thesis #2 here the stumbling block is the definition of the psychological (at least binary) function ‘diverting attention from X to Y .’ But even assuming the existence of a satisfactory account of ‘diverting attention from X to Y ,’ in which X and Y would have to range over properties, it still remains unexplained how the target property Y is an abstract object or a structure, as it is guaranteed by theory #3.

The original definition of abstraction by Dedekind is still unrivalled. It simply says that abstraction is the convention according to which it is always a function that determines which properties an object is going to have, so that we can only grasp the object via our understanding of the function.

Strategic option #2. A commitment to Fregean abstract dualism, whereby we will accept the existence of abstract entities that are not themselves objects. The now classical examples of such entities are Fregean functions and Fregean concepts, as opposed to Fregean courses of values. In this framework, categories and patterns will belong with Fregean functions and concepts, since they are also unsaturated or incomplete and thus are also predicative, like concepts.

Strategic option #3: The Problem of existence. This concept is interpreted as meaning reliable knowledge about the entities being asserted to exist. In the theory of patterns, knowledge is to be recognized as reliable when it meets the requirements of the method of the wide reflective equilibrium. With this method, we arbitrate between competing claims to reliability based on different kinds of evidence, e.g., formal versus informal. As was already suggested, truth is defined in terms of the coherence theory.

III. Varieties of Structuralism

I turn now to my block of open questions and would like to begin by saying a few words about the varieties of structuralism. As it would be desirable to have a minimal definition of structuralism, applicable to all its varieties, I propose to define it informally as a doctrine according to which the object of mathematical thought and the content of a mathematical theory is the presentation of a structure.

We can avoid the obvious circularity of this definition by giving the word 'structure' its more usual meaning, as a domain of objects together with relations and functions defined in it. An example of such a structure is a group, defined as

- (1) A non empty domain G , together with a binary function on G denoted by \circ and
- (2) such that \circ is associative and
- (3) there exists a unique element i which is an identity for \circ satisfying the equalities, for all x in G ,

$$\begin{aligned}x \circ i &= x \\i \circ x &= x\end{aligned}$$

and

- (4) For every element x of G there is a unique inverse y with respect to \circ , satisfying the equalities

$$\begin{aligned}x \circ y &= i \\y \circ x &= i.\end{aligned}$$

If we now analyse this definition we soon meet our familiar objects set (this is the domain), relation and function. This provides a set-theoretical concept of structure and makes it possible to talk about structures as mathematical objects themselves. The structure is then a n -tuple, in my example the ordered pair $\langle G, \circ \rangle$, in which case the thesis that the object of mathematics is the presentation of structures includes the conception of set theory as the canonical language of mathematics.

Charles Parsons believed that a definition of structuralism was threatened by the following dilemma: if we posit structures or patterns as prime objects, we are unfaithful to the structuralist main thesis according to which prime objects are only positions in patterns; but if we accept the set-theoretical framework we have to provide an account of sets that is structuralistically meaningful. In spite of the fact that this dilemma can be met in more than one way, I want to leave this topic by remarking that it cannot be surprising that a definition of structuralism runs the risk of being circular. This is due to the fact that the concept of structuralism stands or falls with the more general concept of function, and for this one, as Frege has shown, there is no noncircular definition.

By now the best known taxonomy of the structuralist family is due to Charles Parsons in his 1990 paper 'The Structuralist View of Mathematical Objects.' There we find a division into eliminative and non-eliminative, which he also calls the metalinguistic form of structuralism.

Eliminative structuralism, associated with the names of Paul Benacerraff, Michael Resnik and Stewart Shapiro, holds that statements about natural numbers are really about any simply infinite sequence or progression. In general, statements about any kind of mathematical objects are general statements about structures of a certain kind and provide thereby the elimination of reference to those mathematical objects. Eliminative structuralism also includes if-thenism or Logicism, first associated with the name of Hilary Putnam, and includes also Modalism, the program of eliminating mathematical objects in favour of modalities.

Non-eliminative structuralism takes the language of mathematics at face value and does not require that the objects referred to be any more objectively determined than the language itself already specifies. Structure is a metalinguistic notion, like a model in which the domain is given by a predicate, and the relations and functions by more predicates and functional terms. In the literature, there is another well-known taxonomy due to Michael Dummett in his *Frege: Philosophy of Mathemat-*

ics. There he divides structuralism, as a philosophical theory about the reference of mathematical terms, into weak and strong.

In weak structuralism, associated with Frege, the definition of a mathematical object includes applications and the structuralist tone is obtained from the fact that anything of suitable structure will be considered. His example is the definition of the real numbers. For Frege, the real numbers apply to any set of relations, producing what he calls a magnitude domain. The conditions for the domain are laid down but the real numbers, i.e., the objects being defined, are defined as extensions covering all such domains.

Dummett divides strong structuralism into hard headed, to be associated with the name of Paul Benacerraff, and the mystical, to be associated with the name of Michael Resnik. Whereas hard-headed structuralists allow structures that also have non-structural properties, mystics only allow structures with structural properties.

I sketch now a short chronology of more than one hundred years of structuralist literature. Structuralism begins at the end of the XIX century (1888) with Dedekind, in particular his essay 'Was sind und was sollen die Zahlen?' I locate the period of early structuralism in the middle of the 20th century, with two of the survivors of the Hilbert program, Paul Bernays in 1950, and Wilhelm Ackermann in 1960. Classical Structuralism begins with Paul Benacerraff in 1965 with his essay 'What Numbers Could Not Be.' Also in 1965, Charles Parsons published his first structuralist work and by the end of the sixties, 1969, Quine's ontological relativity shows a striking kinship with the structuralist approach to reference. Michael Resnik's theory of patterns, first and second version, and Stewart Shapiro's 'Mathematics and Reality' turn up in the 80's and in 1990 Charles Parsons proposed his non-eliminative metalinguistic version of structuralism.

In spite of the fact that classical structuralism began roughly at the time that category theory established itself, the structuralist meaning of this new and related theory was only first hinted at by John Bell in his 1980 lecture 'Category Theory and the Foundations of Mathematics' and then fully characterized in 1993 by Colin Mc Larty in a rival paper to 'What Numbers Could Not Be' with the title 'Numbers Can Be Just What They Have To.' A very perceptive analysis of the dispute as to which of the two concepts, category or set, is more appropriate for foundations, is to be found in J.P. Marquis's 1995 essay 'Category Theory and the Foundations of Mathematics: Philosophical Excavations.'

IV. The Identity Problem

Identity as a problem makes its first appearance in classical structuralism in connexion with Frege's difficulties with identities of the form

$$n = \text{Julius Caesar.}$$

To report on the eliminative position on identity I shall use expressions with the general form

$$n = s$$

in which n is a cardinal number and s a set.

$$\text{Example: } 5 = \{\{\{\{\{0\}\}\}\}\}.$$

Relative to expressions of this form Frege's dual problem was

- (i) To compute always a *truth-value* for such identities and
- (ii) To determine the meaning for the insertion in s of a name or also of a definite description.

We can split such expressions in types, with basically 3 items:

Type 1: An arithmetical term is inserted in the right hand side of the identity.

$$\text{Ex.: } 5 = 2 \cdot 2 + 1.$$

Type 2: A term is inserted.

$$\text{Ex.: } 5 = \text{the number of continents.}$$

Type 3: A proper name is inserted.

$$\text{Ex.: } 5 = \text{Julius Caesar.}$$

Recall that for Frege there is a universe U that contains the denotations of all names and definite descriptions. Thus, it makes sense for Frege to ask of any two names n_1 and n_2 whether they are the names of the same object. Frege complains that from the system of definitions in the *Grundlagen* he cannot decide his problem (ii) for expres-

sions of type 3, so that he could not fix the meaning of identities of the form

$5 = \text{Julius Caesar}$.

The Archimedean point of eliminativism, as far as identity is concerned, consists in denying that one always has to determine meaning in any expression with an occurrence of the binary symbol '='. Instead, it must be possible to consider expressions of type 3 as a-semantic or meaningless, in the same sense that one considers meaningful to say that '5' is the solution of the equation

$$x \cdot 4 = 20$$

and meaningless to say that 'Julius Caesar' is the solution of the same equation.

In eliminativism, an identity is only meaningful if one has a theory in whose language one can formulate and interpret individuation clauses. In general,

$$a = b$$

is meaningful if there is a class of objects K , which contains both a and b and such that the set of definitions which allow the recognition of a and b as being the same K also allows the computation of the truth-value for

$$a = b.$$

If a and b are names of horses, we can decide whether they are the same horse. Individuation parameters are here names of parents, date of birth, place of birth etc. And if a and b are both names of classical gods we can also decide whether they are the same god. But the individuation parameters of a classical god cannot be the same as those of a horse, since what makes an object be a particular horse cannot make it be a particular classical god. In the end, the meaning of the question as to whether two objects a and b are indeed the same object is derived from the implicit assumption that a and b both belong to a certain fixed K . And so what is constitutive of the identity of an object is K dependent, or rather dependent on the theory in which K is used.

It remains to consider the possibility of, instead of determining a meaning for formulas of type 3, to rather compute their truth-value in the sense of Frege's problem (i) above. Since we cannot have identities across theories, formulas of type 3 can only be false.

But as it is not easier to see that they are false rather than meaningless, in eliminativism one argues that they are either false or meaningless.

I would like to sketch now the identity problem in the theory of patterns and for that purpose I begin by explaining the basic vocabulary.

A pattern P is a composite entity defined by a relation R over the objects of P , called the points of the pattern. Patterns are also called structures and points are also called positions. The notation for a pattern is the notation for a structure, as already used. Thus

$$\langle N, s \rangle$$

denotes a pattern, whose points are the natural numbers and whose relation is the successor function. A pattern is not a set and differs from a set in that the pattern is not defined via its points, as a set is defined via its elements, but only by the relation R . The objects of P , the points or positions, have no identifying properties beyond the ones imposed by the relation R and so no extra structural property defines identity.

The paradigm instance of a point or a position is actually the point of the plane. If you consider the unit circle with centre c and diameter $a b$, the three points a, b, c considered in isolation have no individuating properties, but as elements of the circle they can be distinguished. In a pattern, the only individuating properties of the objects are created by the relation, as it is the case with points in geometry, like the point c that is the centre of our unit circle above.

Patterns can be in two kinds of relations:

- (i) The first kind is the relation from pattern to instance, called *instantiation*, and
- (ii) The second kind is the relation from pattern to pattern.

Of this second kind the following, ordered from weak to strong, are relevant:

Occurrence,
Mutual occurrence,

Inclusion,
Equivalence,
Congruence.

From these the strongest is congruence, which is the relation that obtains between two patterns P_1 and P_2 which have isomorphic models or instances. I call two models isomorphic if there is a structure preserving function with a structure preserving inverse.

We now turn to Occurrence, which is the relation that obtains between two patterns P_1 and P_2 when P_1 is isomorphic to a pattern definable in P_2 .

As an example, we can take P_1 to be $\langle N, s \rangle$ and P_2 to be $\langle Q, +, x \rangle$.

A pattern P_1 is a subpattern of a pattern P_2 if every point in P_1 is in P_2 and P_1 occurs in P_2 .

Occurrence is reflexive and transitive and thus is a quasi order.

If you now consider the following two patterns,

$$\begin{aligned} P_1 &= \langle N, s \rangle \\ P_2 &= \langle N, < \rangle \end{aligned}$$

it is obvious that every point in P_1 is in P_2 and that we can define 'successor' in P_2 , by saying that

$$\begin{aligned} n \text{ is the successor of } m \text{ if and only if} \\ m < n \text{ and there is no } t \text{ such that } m < t < n. \end{aligned}$$

Since identity is itself an equivalence relation it is natural to search in the list for equivalence relations that could be seen as a proxy identity for patterns. The equivalence relations are: mutual occurrence, equivalence and congruence. As this last one is the strongest of the three, it suffices to show that identity is not equivalent to congruence.

It is enough to show that congruence cannot be a sufficient condition for identity between patterns. As a reductio argument, let us assume that it is. Then as the N -progression and the $2 \times N$ progression are congruent, by hypothesis they are identical. Therefore, their points are identical and we then have equalities like $1 = 2$. So congruence cannot be a sufficient condition for identity.

We could try to promote the use of the relation Occurrence as a definition of identity by interpreting occurrence as interpretability of theories, using definitions.

We say that a theory T_1 is interpretable in a theory T_2 just in case there is a sequence of definitions of the primitive symbols of T_1 in T_2 , which we denote by

$$\{D T_1 T_2\}$$

which added to T_2 produces the theorems of T_1 as theorems of the new theory

$$T_2 + \{D T_1 T_2\}.$$

We work this out with an example.

We assume two second-order theories N' and $N+$ with the following primitive symbols:

- (i) N' has 0 and monadic ' $'$

and

- (ii) $N+$ has 1 and dyadic '+.

The axioms of N' are

- (i) 0 is not a successor;
 (ii) Successor is a 1-1 function;
 (iii) Second-order induction.

The axioms of $N+$ are the same, replacing the monadic successor function by ' $+ 1$ ' and defining 0 as the number whose successor is 1. An so we have both:

- (i) N' interpretable in $N+$.

The definition needed is $x' = x + 1$.

- (ii) $N+$ is interpretable in N' .

Two definitions will do, one in which 1 is defined as the successor of 0 and the other an inductive definition of $a + b = c$.

This will show that every model of N' has an occurrence of $N+$ inside and that this relation is symmetrical. So both theories are recip-

roccally interpretable. However, they cannot be regarded as being the same, because N' is complete and categorical and N^+ is neither complete nor categorical.

I sketch now the outline of the problem concerning the restriction of identity to a relation within a pattern. To carry out the ongoing analogy with geometry, we start by considering the possibility of developing geometry not as a theory of space but rather as a cluster of theories about configurations in some space.

In this framework, the points are now the positions within the configurations but not in the including space. We then define identity relative to objects of the configuration. In the ideal case there will be a separate theory for each configuration and in this case each theory will have an absolute identity predicate, ranging over all the elements of its universe.

If we apply now this analogy to arithmetic, we can certainly say that it deals with a certain configuration, and the identity of numbers can be taken as absolute in the sense that for any x and y in N either $x = y$ or $x \neq y$.

There will be a slight ambiguity in the notation of relativized identity symbols. We can introduce the name of the theory or of the configuration as a subscript in $=$. If we now interpret the elements of the N progression $0, 1, 2, 3, \dots$ to be positions, we can make sense of a statement $2 \neq_n 1$ as meaning that they are different positions of the N progression; and about the progression $0, 2, 4, 6, \dots$ we may want to express that 2 has in the e-progression the position that 1 has in the n progression by writing $2 =_e 1$. But then within the same pattern one will have formulas like $4 =_e 2$, to express that 4 has in the e-progression the position that 2 has in the n -progression, but we also have $4 \neq_e 2$ to express that 4 and 2 are different points of the same e-progression.

It would appear to have been more natural to try to define identity between patterns by a formula like

P_1 is the same as P_2 if and only if
 for every x , x is a position of P_1 if and only if
 x is a position of P_2 and the relations of P_1 are extensionally identical to the relations of P_2 .

Such a formulation would however have the disadvantage of a commitment to admitting positions of different patterns into a single universe of discourse.

Therefore, patterns and positions in different patterns are not in the field of the identity relation. In conclusion, what are then the options for cross pattern relations? There are at least three possible options.

The first is to accept a many-sorted framework in which each pattern would have its own separate universe of discourse. The second would be exploring again the analogy with geometry and positing a space of positions or points, from which patterns could be constructed. Identity in this framework would then apply to all points of this space, but then of course the theorems about this space would not deal with individual points.

In principle, there is a third and last possibility, which is with the framework of set theory. The points of a pattern would be conceived as sets and the theory of patterns would be reduced to set theory. The obvious disadvantage of this proposal is to force the interpretation of patterns as sets, which is something that they are not.

What does the rejection of congruence as the identity relation for patterns amount to? It amounts to the following. If two patterns are congruent then they are isomorphic, and isomorphic objects share all the same properties. But by Leibniz's conception of identity we cannot have distinct objects with no difference between them. So if those two patterns are different, but yet indiscernible, they violate Leibniz's identity law. Shall we then interpret this as meaning that Leibniz's law of identity is to be rejected or as meaning that the congruence relation has to count as identity?

It turns out that Leibniz's identity law is not the only conception of identity available. Less known is the so called Behmann's identity, also followed by Ramsey and the early Wittgenstein, according to which two objects can be different and yet satisfy exactly the same properties. Behmann's identity would then make justice to the theory of patterns by allowing two patterns to be different and yet be isomorphic or congruent.

There is a simple logical relation between Leibniz's identity and Behmann's identity and that is that a formula expressing Behmann's identity is the dual of a formula expressing Leibniz's identity. In this case one could say that pattern congruence is the dual form of identity. I close this topic by remarking that even this solution leaves an after taste of an ad hoc expedient because both conceptions of identity are supposed to handle identity between objects. But since patterns are not objects, we should not expect any known form of identity to hold for them.

V. From Patterns to Categories

In his paper on Cantor's Continuum Problem, Gödel underlines the shift in conceptual scheme that occurred with the introduction of Ellmann and McLane's theory of categories in relation to the iterative conception of set. I would like to say something about this shift and sketch the outlook of the new conceptual scheme before comparing the structuralist with the categorical point of view.

To explain the notation: if g is a function from a set A to a set B and f a function from the set B to a set C , then by

$$f \circ g$$

I mean the function h from A to C such that

$$H(x) = f(g(x)).$$

We call $f \circ g$ the *composition* of both functions.

This notation is equivalent to $f \circ g(x)$. If we define an isomorphism between two sets A and B to be a function f from A to B with an inverse f^{-1} from B to A , then the composite

$$f^{-1} \circ f$$

is the identity function on A , and

$$f \circ f^{-1}$$

is the identity function on B .

As far as terminology is concerned, a relation in a category will be called usually a morphism, but the notion matches in many examples the profile of a function in the usual sense. The simplest informal description of a category C is as a totality O , which we posit to contain the objects of C . We characterize the totality O by the following existence statements:

- (i) For every pair A, B of elements of O , there is a totality $M(A, B)$, called the morphisms from A to B in C ;
- (ii) For every triple of elements A, B, C in O there is an operation from pairs of morphisms in

$$M(A, B) \times M(B, C)$$

to morphisms in $M(A, C)$ called the composition of morphisms in C . In the notation above, if f is a morphism from A to B and g a morphism from B to C then

$$g \circ f$$

is the morphism from A to C , their composition;

- (iii) For every element A in O there is in $M(A, A)$ a morphism $\text{Id } A$, called the identity on A .

From morphisms two things are in turn required:

- (1) The associativity of their composition, and
- (2) left and right identities.

Thus if f is a morphism from A to B and g a morphism from B to C and h a morphism from C to D , then we want the equality

$$h \circ (g \circ f) = (h \circ g) \circ f$$

to be true.

The left and right identities are captured in the two formulas

$$(\text{Id}_B \circ f) = f$$

and

$$(f \circ \text{Id}_A) = f,$$

when f is a morphism from A to B . The best-known examples of mathematical structures are examples of categories. The philosophically interesting feature of this concept is that what distinguishes a category is the category's own morphisms, not the category's own elements. In a category's instance these elements are structured sets with structure preserving maps. However, a category is anything that fulfils the definition above. For my purpose, I use abstract sets as an example of a category. Whereas concrete sets collect concrete individuals together, an abstract set is an idealization of a concrete set

with the significant difference that it is described by its functions rather than by its concrete elements. So if A and B are abstract sets and f a function from A to B then any function g from B to C has a composite with f , $g \circ f$ from A to C . Notice that we do not define composition by an action on the elements of A . We can now have axioms to stipulate the associativity of composition and for the existence of the identity functions. In particular, every set A has a function Id_A from A to A such that the following identities are true:

$$\begin{aligned} f \circ \text{Id}_A &= f \text{ and} \\ \text{Id}_A \circ h &= h, \end{aligned}$$

when f is a function from A to B and h is a function from C to A . Thus abstract sets and functions are an instance of a category.

At this point, it is useful to introduce an axiom guaranteeing the existence of the abstract singleton. This axiom asserts the existence of a set, denoted by 1 , with the following property: no set can have more than one function with range 1 . It is then possible to define abstract element as being actually also a function. Thus an element x of an abstract set A is any function x from 1 to A , and as such elements have also only structural properties.

This shift from object (or element) to function is the essence of the new categorical outlook, first described by Lawvere in 1964. In categorical set theory sets and functions have only structural properties. And as I said above we have two ways to interpret this result: either so that it would seem that the structuralist program is thereby fulfilled or so that it would seem that the structuralist program is thereby made redundant.

I take the easiest disjunct first, the one about redundancy. To define a theory as redundant relative to another it has to be assumed that they are both theories of the same kind pursuing the same goals. Now, this is not the case in the present debate, since structuralism does not offer a mathematical theory that can compete with categorical set theory. In this sense, structuralism is not redundant and it has to be justified on its own merits, as any standard philosophical theory.

It remains to settle the question of whether categorical set theory fulfils the structuralist program. Here the whole question rotates around the interpretation of the concept of fulfilling a program. There was obviously no structuralist program around when categories and abstract sets appeared and produced the shift that Gödel mentions

in his article. So, in this interpretation there is no fulfilment of the structuralist program. But there is another interpretation possible, namely as viewing the philosophical theory as an expedient rhetorical device that captures the spirit and the achievements of categories and categorical set theory.

M. S. Lourenço
University of Lisbon

A refutação da teoria empirista da consciência*

M. S. Lourenço
Universidade de Lisboa

1. A concepção empirista da *res cogitans*

A finalidade deste capítulo é proporcionar uma descrição dos princípios mais gerais que são utilizados nas discussões de Wittgenstein acerca da natureza da experiência dita interior¹. Dois géneros principais de argumentos devem ser em *in limine* separados:

- i) Argumentos acerca da possibilidade de determinar se a natureza da experiência interior é privada, num sentido da palavra «privado» a ser definido oportunamente; e
- ii) Argumentos acerca da possibilidade de determinar a natureza do meu conhecimento da consciência de outrem; em particular, se o meu conhecimento da consciência de outrem é para ser considerado como sendo apenas indutivo ou analógico.

Ambas² as questões caem dentro do âmbito da questão mais geral que consiste em determinar que género de relação existe dentre as experiências ditas interiores e a sua representação linguística ou, melhor ainda, que género de relação existe entre as expressões da linguagem

* Capítulo 2 de *Espontaneidade da Razão*, de M. S. Lourenço (Lisboa: INCM, 1987).

¹ Acerca do problema da caracterização geral da natureza em *Philosophische Untersuchungen* e *Zettel* ver Malcolm, Norman «Wittgenstein on the Nature of Mind» in *Studies in the Theory of Knowledge* (1970), pp. 9-29. O emprego do termo *consciência* no texto é para ser compreendido exactamente como o termo alemão *Bewußtsein*.

² Destes dois géneros de argumento de carácter geral seguem-se dois corolários principais, um acerca das relações entre as linguagens L_p e L e outro acerca da possibilidade de eliminar prefixos como «eu sei» em contextos como «eu sei que X » quando X é uma descrição de um estado de consciência.

natural que se referem a experiências interiores e estas mesmas experiências. Exemplos conhecidos destas experiências são, do ponto de vista perceptivo, as sensações e as percepções, do ponto de vista afectivo as emoções, do ponto de vista volitivo as intenções, do ponto de vista cognitivo os conceitos e os juízos.

Como já foi observado no capítulo anterior, um aspecto crucial da nova orientação filosófica de Wittgenstein é a sua rejeição do que então chamámos *essencialismo*, uma concepção segundo a qual é possível determinar o sentido de uma expressão acerca de um objecto *S* através da determinação da essência de *S*. Assim, à luz da rejeição do essencialismo já descrita, não nos esforçaremos por determinar o género de relação que existe entre as expressões da linguagem natural que se referem a experiências interiores e estas mesmas experiências através da determinação da sua essência. De acordo com os resultados introduzidos no capítulo anterior, não procuraremos adivinhar qual é o sentido de uma expressão acerca de uma sensação mas, antes, procuraremos seguir o modelo proposto por Wittgenstein, nomeadamente o de considerar uma sensação em particular, no seu exemplo a sensação de dor. Tentaremos assim, através de uma discussão do que é a determinação do sentido de expressões que se referem à experiência interior de dor, ganhar a possibilidade de compreender a relação que existe entre a experiência interior e a sua representação linguística. O nosso primeiro objectivo é agora o de delinear a concepção pressuposta nos argumentos de teor empirista acerca da natureza da existência interior e da sua relação com os meios linguísticos que a representam.

Segundo a concepção a delinear, a minha experiência interior — e portanto as minhas sensações, percepções, emoções e intenções — só a mim próprio diz respeito e é, neste sentido, uma experiência puramente privada. Não é de todo possível que uma outra pessoa tenha acesso à minha consciência, e aí procure ter as experiências que eu neste momento estou a ter. No caso do exemplo proposto por Wittgenstein, só me é possível determinar o que se pode entender por uma dor quando eu observo, à luz de minha própria experiência interior de dor, i.e., das minhas sensações de dor, o que é que constitui para mim a experiência de dor. Logo, o método que se utiliza para chegar à formulação do conceito geral de uma dor é o método indutivo que consiste em observar em cada caso o que é que constitui, para o caso sob consideração, a experiência de dor e daí inferir para o caso geral, sendo assim a experiência privada que constitui o fundamento

do conceito geral. Ao tentar compreender o sentido da palavra «dor» não tenho mais a fazer do que me *recordar* do que é minha experiência de dor, e assim o sentido da palavra «dor» é a experiência interior de dor que se observa na anamnese. Quando a seguir me pergunto como me será possível obter uma compreensão do que é a experiência de dor de outrem, por exemplo a experiência de dor de Job, a resposta é que simplesmente se torna suficiente transferir para a experiência interior de Job aquilo que constitui a minha própria experiência interior e portanto, pela utilização desse método indirecto, chego a poder determinar em que é que consiste a experiência de dor de Job.

Esta transferência do meu caso privado e pessoal para o caso de Job pode ser entendida, por sua vez, sob duas formas. Na primeira destas formas trata-se do conhecimento da experiência de Job, como uma experiência de dor, só me ser acessível através do processo indirecto, nomeadamente o da inferência analógica já referida. Na segunda das formas de transferência trata-se de que a designação da experiência de Job como sendo uma experiência de dor é indirecta agora no sentido que, ao ser aplicada não ao meu próprio caso mas ao caso de Job, só o posso fazer através da observação dos sintomas periféricos uma vez que, *ex hypothesi*³, as experiências interiores de Job, as suas experiências de dor, não me podem ser imediatamente dadas na minha consciência. Finalmente resta observar que ainda que a transferência da minha compreensão do sentido da palavra «dor», através da consideração da minha experiência de dor, fosse suficiente para entender a compreensão da experiência de dor e, correspondentemente, do sentido que para Job tem a palavra «dor», a minha própria experiência difere ainda essencialmente da experiência de Job. Em primeiro lugar, como já foi observado no capítulo anterior, segundo a teoria que delineámos, a palavra «dor» só adquire o seu sentido através de sua conexão com as imagens mentais na minha própria consciência, das minhas experiências privadas de dor e, em segundo lugar, do ponto de vista epistemológico, a evidência da minha percepção de mim próprio leva-me a concluir que só acerca de mim próprio é que posso ter conhecimento acerca da possibilidade de ter ou não uma dor, enquanto acerca de Job só posso ter uma crença acerca da possibilidade de Job ter ou não uma dor que diz que tem,

³ A hipótese que serve de postulado ao argumento empirista consiste na proposição de que só existe acesso ao que é imediatamente dado na consciência.

uma vez que no seu caso não consigo ter acesso à sua dor mas apenas às suas expressões de dor, i.e., aos seus sentimentos periféricos, estando assim sujeito às diversas possibilidades de erro na interpretação dos seus sintomas.

Terminamos aqui a delineação das concepções normalmente aduzidas a favor da caracterização da natureza da experiência interior como sendo essencialmente privada de tal modo que o meu conhecimento da minha própria consciência é o único fundamento para a compreensão do que é a consciência de outrem. Da consciência de outrem não chego nunca pois a ter verdadeiramente conhecimento. A estrutura e subtileza dos argumentos de Wittgenstein contra a concepção delineada será posta em evidência através da sua disposição gradual. Essencialmente, trata-se, segundo Wittgenstein, nos argumentos aduzidos a favor da concepção exposta da consciência, de uma análise adequada do que é a natureza da linguagem e, em particular, do emprego de expressões cuja denotação são sensações.

2. A demonstração da inexistência da linguagem L_p

Começaremos agora a exposição da doutrina de Wittgenstein acerca da natureza da consciência pelo exame da hipótese, já delineada na secção anterior, segundo a qual a única fundamentação do conceito geral de dor reside na minha experiência privada de dor. Aproveitamos agora a ocasião para observar que esta hipótese, associada na secção anterior apenas com a filosofia idealista, tem sido de facto usada, explícita ou implicitamente, no período da Filosofia Moderna, por outros filósofos além dos proponentes da doutrina idealista. Obviamente a hipótese aparece em Descartes, uma vez que, na quinta das suas *Meditações Metafísicas*, Descartes defende a ideia de que o conhecimento certo só se encontra representado no conhecimento que eu tenho da minha própria consciência, em comparação com a qual o mundo dito exterior e portanto a consciência de outrem me aparece como sendo sujeita a dúvida. A hipótese da natureza essencialmente privada da consciência e, em particular, de uma correspondente linguagem privada, é dominante na filosofia de Locke. Com efeito, segundo Locke, no meu diálogo com o leitor desta página, cada palavra que eu escrevo, por exemplo a palavra «meditação», só é compreensível para o leitor através de uma associação, que o leitor realiza na sua consciência, cujo conteúdo é a imagem mental *causada*

pela palavra escrita por mim. Esta, por sua vez, é apenas para ser considerada como um sinal da imagem mental que *eu* associo à palavra «meditação». Assim, a palavra «meditação» constitui uma imagem do que é a minha imagem mental não sendo por isso possível ao leitor desta página fazer a mesma associação que eu faço e chegar assim à mesma imagem mental. A concepção de Locke sobrevive no empirismo contemporâneo, em particular na sua formulação fenomenalista. Nesta formulação procura-se representar os objectos do mundo exterior apenas como construções efectuadas a partir dos dados dos sentidos do sujeito cognoscente. As asserções que se fazem acerca dos objectos do mundo exterior têm que ser reescritas, à luz da concepção fenomenalista, sob a forma de asserções acerca dos dados dos sentidos do sujeito cognoscente. Estes dados do sentido são de natureza puramente privada, no sentido em que nenhum outro sujeito cognoscente pode ter os dados dos sentidos que eu tenho em cada momento dado. De um modo geral, pois, aproveitamos a ocasião para generalizar a adopção da hipótese a todas as concepções que se referem ao problema do fundamento do nosso conhecimento acerca do mundo exterior, e das diversas análises propostas para legitimar epistemologicamente asserções acerca da existência do mundo exterior e da consciência de outrem⁴.

Como o que é peculiar e inovador na posição de Wittgenstein não é tanto o facto de se opor à orientação predominante da filosofia moderna mas antes a sua descoberta de que o que é essencial a esta orientação é a concepção da existência de uma linguagem essencialmente privada, esforçar-nos-emos agora por definir em que condições é que a linguagem pode ser considerada essencialmente privada⁵. Para alcançar esta definição convém começar por observar que o que temos em mente definir é uma linguagem que é privada mas não no sentido contingente de ser uma linguagem falada por um único indivíduo, em contraste com uma linguagem falada por uma comunidade de indivíduos, ou ainda de ser uma linguagem contingentemente

⁴ Na filosofia da matemática de inspiração idealista, está concepção é elegantemente defendida por Brouwer na sua comunicação «Consciousness, Philosophy and Mathematics» in *International Congress on Philosophy X*, Amesterdão, 1948.

⁵ Para a definição da linguagem L_p é útil ter em mente o ensaio de Castañeda, Hector-Neri, «The Private Language Problem» in *The Encyclopedia of Philosophy*, org. Paul Edwards, Nova Iorque, pp. 458-464.

privada no sentido de ser a linguagem secreta de uma única pessoa que possui além desta, a sua linguagem pública⁶.

A minha linguagem é privada, para Wittgenstein, no sentido essencial em que não é possível a outra pessoa, além de mim próprio, compreender o sentido das expressões que eu emprego. A razão pela qual é impossível que outra pessoa compreenda a minha linguagem privada é porque as expressões de minha linguagem privada têm como denotação as minhas experiências interiores, às quais só eu posso ter acesso na minha consciência. Neste sentido essencial a minha linguagem privada implica de facto dois géneros de privacidade, nomeadamente a privacidade de posse, no sentido em que as experiências interiores denotadas na minha linguagem só podem ser minhas e ainda a privacidade epistemológica, no sentido em que só eu posso ter conhecimento do que são as experiências interiores a que a minha linguagem privada se refere.

Em particular, alcançamos uma ideia da estrutura presumível de uma tal linguagem considerando o que seria nesta linguagem uma proposição elementar. Para a construção destas proposições o que é essencial é dispor dos nomes de que as proposições elementares são compostas e estes, por sua vez, poderão ser formados através da concentração da atenção numa dada sensação S , à qual é a seguir atribuída um nome, digamos « S » e finalmente fixar na minha memória a conexão entre o nome ou a designação da sensação e a própria sensação. A partir destes nomes, assim definidos, chego à possibilidade de, através da iteração destas operações privadas de definição, construir a minha linguagem privada. Trata-se pois essencialmente de que me é possível estabelecer uma conexão entre uma palavra P da minha linguagem privada L_p e uma sensação S , uma conexão que fixo na minha memória. O corolário óbvio deste processo de construção da minha linguagem privada L_p é que o nome « S » será ainda, no futuro, o nome da sensação S sempre que esta sensação se oferece à minha atenção⁷. Finalmente, podemos concluir esta definição de «privado» com a aposição das duas observações que se seguem. Em primeiro

⁶ Ver a análise desta pressuposição e seu uso em argumentos recentes contra a posição de Wittgenstein aqui descrita em Hacker, P. M. S., «Private Language and Public Speakers» in *Insight and Illusion*, Clarendon Press, Oxford, 1972, p. 215 e segs.

⁷ Assim a linguagem privada L_p consiste em definições D_1, \dots, D_n de nomes de sensações $D_i(S)$ tais que, a definição $D_i(S)$ num dado momento t é condição necessária e suficiente da definição $D_{i \cup k}(S)$ num momento $t \cup k$, com $t < k$.

lugar tem-se, como corolário da definição de «privado», que não se trata de determinar se seria possível a Robinson Crusoe desenvolver uma linguagem de que ele seria o único utente, uma vez que Robinson Crusoe, embora só na ilha, poderia desenvolver uma linguagem que fosse susceptível de ser compreendida por outrem, ao contrário da linguagem L_p . Em segundo lugar, o sentido de «possível» (ou impossível) usado nestas discussões é o sentido forte de possível, nomeadamente *logicamente possível (ou impossível)*. Neste sentido o conceito da existência de uma linguagem privada é impossível por permitir a derivação de uma contradição. Exemplos do mesmo género de impossibilidade são os conceitos de, e.g., «objecto azul inextenso» ou «esferas de 5 faces».

Passamos agora à exposição dos argumentos de Wittgenstein contra a hipótese da existência de uma linguagem privada no sentido de «privado» que acabamos de precisar. Wittgenstein começa por observar que o meu último passo na construção de nomes de L_p , nomeadamente a sua fixação na minha memória, significa de facto que no futuro terei de que ter a recordação certa da conexão entre o nome « S » e a sensação S . Ora uma recordação da conexão entre o nome « S » e a sensação S só pode ser considerada como sendo a recordação certa quando o conteúdo da recordação está de acordo com o meu segundo passo na construção de « S », nomeadamente que « S » é de facto o nome de S . Portanto estamos agora diante da questão de saber como é que no futuro se poderá decidir se o meu emprego de « S » como nome de S é de facto justificado; por outras palavras, precisamos de encontrar um critério que nos permita separar aqueles empregos de « S », que estão de acordo com o segundo passo já referido, daqueles que não estão.

No parágrafo 258 de *Philosophische Untersuchungen* recusa Wittgenstein a ideia de que este critério me pode ser dado através da consideração daquilo que no futuro me aparecer como sendo a recordação certa da construção de « S » uma vez que aquilo que no futuro *me aparece* como sendo a recordação certa não é ainda critério para decidir se a recordação é ou não a recordação certa. Ora, por outro lado, a hipótese da possibilidade da construção privada de « S » implica a existência de um tal critério. Assim ao demonstrarmos a inexistência do critério refutamos a possibilidade da construção privada de « S ».

Para compreender a natureza do argumento de Wittgenstein convém ter em mente uma sinóptica da construção de « S » como protótipo da construção de L_p :

- i) A concentração da atenção em S ;
- ii) A atribuição a S do seu nome;
- iii) A fixação na memória de que « S » é o nome da sensação S .

Deste modo, as regras que são usadas na construção de « S » e de L_p , são o que se pode designar, neste contexto, por regras privadas, no sentido de privacidade de posse e de privacidade epistemológica e assim as regras só podem ser possuídas por mim e só eu posso saber como as regras são usadas de facto nas diversas construções que levam à formação de L_p . Mas tal como a minha impressão de que o meu uso do nome « S » é o uso do nome certo de S não chega para fazer desta impressão um critério, assim também a noção de uma regra privada, no sentido especificado, não é suficientemente clara para a poder distinguir de uma impressão, i.e., não posso inferir a existência de uma regra do facto de me julgar a seguir uma regra. Quando me encontro de baixo da impressão, digamos durante a construção de « S », de seguir uma regra para poder prosseguir a construção, não tenho ainda suficiente evidência para me permitir deduzir a existência de uma regra. Para me ser me permitida esta inferência tenho que ser capaz de encontrar, para a minha impressão, uma justificação independente dela que tenha o poder de lhe conferir validade intersubjectiva. Na construção de « S » no entanto a justificação da validade dos passos permanece apenas subjectiva, uma vez que a minha impressão do emprego de « S » ser o emprego do nome certo de S só pode, por sua vez, ser justificado por uma outra impressão. Este processo de justificação de uma impressão através do apelo a uma outra impressão conduz à falácia conhecida pela designação de *regressus ad infinitum*⁸. Logo uma regra privada, e portanto a impressão de seguir uma regra, não é uma regra. Ora como vimos no capítulo anterior uma linguagem ou um jogo de linguagem é definido pelas suas regras. Assim uma «linguagem privada» é uma impossibilidade lógica.

Terminamos esta secção fazendo referência a uma estratégia a propor como um meio de escapar à *reductio ad absurdum* de Wittgenstein e por último à refutação desta estratégia com os termos da *reductio* que acabamos de expor. A estratégia proposta consiste em tentar reganhar para o conceito de seguir uma regra privada o estatuto que este con-

⁸ Uma formulação extremamente elegante desta interpretação do argumento de Wittgenstein aqui proposta é a de Hintikka, Jakko, na sua nota «Wittgenstein on Private Language: Some Sources of Misunderstanding» in *Mind*, 78, 1969, pp. 423-425.

ceito perdeu com o argumento de Wittgenstein, uma vez que se o conceito puder ser reintroduzido a possibilidade de construção de uma linguagem como L_p fica de novo assegurada.

Assim, começamos por admitir que não me é possível nunca saber se sigo as regras de minha linguagem L_p de facto ou se apenas me imagino que as sigo, dado que não me é possível construir uma demonstração de que as sigo efectivamente. Mas em vez de construir uma demonstração de que sigo as regras de uma linguagem L_p efectivamente, posso mostrar que as sigo ao mostrar que me recordo do sentido certo do nome «S» ou, por outras palavras, ao mostrar que tenho dele a recordação certa. Esta acomodação é perfeitamente razoável e uma vez concedida é suficiente para reintroduzir a noção de uma regra privada e com ela a possibilidade de construir L_p .

A refutação desta proposta consiste em observar que a única razão pela qual eu julgo poder associar com o conceito de seguir uma regra privada um conteúdo fixo é o facto de eu associar com ele não este conteúdo mas imagens mentais de diversas formas que aparentam dar ao conceito o conteúdo que carece. Mas, como já foi estabelecido no Capítulo 1 de *Espontaneidade da Razão*, na formulação do sentido de uma expressão como «S» em termos das imagens mentais com ela associadas, estas imagens por sua vez carecem tanto de sentido como as expressões que elas supostamente são introduzidas para explicar. Em particular, não está ainda de todo determinado como é que estas imagens por sua vez se usam, i.e., que critérios se adoptam para regular o seu uso de modo que a questão não é resolvida mas apenas deslocada⁹. Finalmente resta observar que quando eu digo que posso encontrar a minha recordação do que foi a sensação S e do seu nome «S», podendo assim justificar o meu emprego de «S» em circunstâncias futuras, tenho também que incluir que a recordação em questão tem que ser a recordação certa visto que nenhuma outra serve para realizar a justificação desejada. Mas, como já foi demonstrado durante a *reductio ad absurdum*, não estamos em condições de poder especificar um critério que nos permita analisar estas recordações de modo a poder decidir quando é que uma recordação é a recordação certa. Os critérios que apelam para as minhas impressões não servem para o processo de decisão e assim, na ausência de qualquer critério, deixa

⁹ Acerca do papel de *Vorstellungen*, no texto referido pelo termo *imagens mentais*, ver a análoga refutação do seu uso em «Widerlegung des Idealismus» in *Kritik der Reinen Vernunft*, Elementarlehre, II T. 1 Ab. II Buch 2 Hauptst.

de me ser possível falar de minha recordação de «*S*» como sendo a recordação certa.

Uma variante elegante desta estratégia consiste em argumentar que embora seja possível que eu não consiga oferecer um critério que justifique uma decisão acerca de eu ter ou não seguido uma regra privada, pode no entanto uma terceira pessoa estabelecer se eu segui ou não uma regra, nomeadamente verificando se o meu emprego do nome «*S*» é determinado pela ocorrência da mesma sensação *S*.

A refutação desta variante consiste na análise da noção de «mesma sensação» ou, o que é o mesmo, na análise do critério de identidade a ser usado a propósito de qualquer par de sensações considerado. De novo temos que fazer apelo às noções introduzidas no Capítulo 1 de *Espontaneidade da Razão*, em particular, à noção de que o sentido de um conceito, como «identidade entre sensações» não é introduzido uma vez por todas mas, ao contrário, é estabelecido gradualmente através das diversas estipulações acerca de seu emprego.

Vale a pena considerar neste momento um exemplo conhecido de Wittgenstein, usado normalmente em discussões acerca da natureza da necessidade lógica e à qual voltaremos com mais pormenor no decorrer do Capítulo 3 de *Espontaneidade da Razão*. O exemplo revolve à volta da noção da aquisição do conceito de sucessor de um número inteiro n , onde se supõe que uma pessoa compreendeu o método de determinar o sucessor de um número inteiro n , nomeadamente ao ter compreendido o sentido de « $n + 1$, para qualquer n ». Esta suposição é a seguir confirmada pelo facto de em diversas ocasiões a pessoa em questão formar sempre correctamente o sucessor do número inteiro dado. No exemplo de Wittgenstein passamos agora a considerar o facto surpreendente que esta compreensão do método de formar o sucessor de um número inteiro n só é satisfeita quando o número é menor do que, por exemplo, 1 000. Quando n é maior do que 1 000 já não é derivável de « $n + 1$, para qualquer n ». Em particular uma variedade de métodos parece ser agora o caso. Assim, quando n é maior do que 1 000 e menor do que 2 000, a formação do sucessor é agora $n + 2$, quando n é maior do que 2 000 e menor do que 3 000, o sucessor de n é calculado como sendo $n + 3$.

Segundo Wittgenstein este comportamento não é irracional e é até justificável uma vez que não existe nada na primitiva instrução do método de calcular o sucessor de um número inteiro n que exclua a possibilidade de se proceder de acordo com a variedade de métodos a que o exemplo se refere. Em particular, a situação não é sequer

esclarecida através de uma nova ordem, novamente a ordem de proceder da mesma maneira, i.e., da maneira usada no cálculo do sucessor de n para n menor do que 1 000, uma vez que é sempre possível responder-se que se está ainda a proceder da mesma maneira. Chegamos assim a compreender a noção proposta por Wittgenstein segundo a qual se deve reconhecer que o conceito de «identidade entre dois métodos» não é definido de uma vez por todas mas tem que ser estabelecido gradualmente e que, em circunstâncias diferentes, serão propostos critérios diferentes. No caso de uma regra será então também gradualmente determinado se a regra foi ou não seguida da mesma maneira e a determinação do facto de se ter seguido a regra correctamente reduz-se à determinação do procedimento estar ou não de acordo com o uso comum. Vale a pena reparar no facto de as circunstâncias do exemplo não serem normalmente encontradas na experiência, serem de natureza puramente contingente e não necessárias uma vez que, como já foi dito, a primitiva instrução não exclui necessariamente a possibilidade de se adoptar a variedade de métodos.

A resistência que normalmente se oferece à adopção da crítica implícita no exemplo de Wittgenstein é causada pelo facto de nos julgarmos de posse de uma ideia clara e inequívoca de identidade, cujo paradigma é normalmente expresso pela primeira das leis de Leibniz, i.e., o axioma da identidade segundo o qual qualquer objecto é idêntico a si próprio. Mas, como é conhecido, para definirmos o conceito de identidade este axioma tem que ser suplementado por outro que especifique como condição necessária de identidade, a possibilidade de se atribuírem os mesmos predicados a objectos considerados idênticos. No caso da nossa discussão da existência possível de uma linguagem privada podemos agora formular a tese crucial, em parte já obtida no Capítulo 1 de *Espontaneidade da Razão*, segundo a qual seguir uma regra é em si um costume. Só onde se verifica a existência de um uso comum e que obrigue igualmente todos os utentes de uma expressão se pode falar no estabelecimento de um critério para o emprego correcto da expressão, e portanto da formação de uma regra que controle esse emprego. Na circunstância de eu me dizer a mim próprio que na minha linguagem privada sempre que ocorre de novo a sensação S eu usarei o nome « S » não tenho ainda uma regra para o uso correcto de « S », uma vez que eu posso fazer com o nome « S » o que eu quiser sem necessariamente contradizer a minha determinação uma vez que, na ausência de uma instância independente de mim próprio, eu posso sempre chamar mais tarde a mesma sensação ao que

eu quiser. Deste modo chegamos assim a um resultado análogo ao da nossa análise da determinação da recordação correcta, nomeadamente de que se eu posso chamar «o mesmo» àquilo que eu quiser, o uso desta noção torna-se por esse mesmo facto irrelevante e tem que por isso considerar-se a noção assim definida como vazia¹⁰.

3. A refutação do solipsismo

O nosso propósito nesta secção é analisar o género de argumentos que, partindo da premissa da existência de uma linguagem privada como L_p , conduzem à conclusão da minha impossibilidade de acesso aos estados mentais de outrem. Com efeito, se eu garanto a existência da minha linguagem privada L_p , garanto igualmente, pelo mesmo raciocínio a existência da linguagem privada L'_p de Job. A consequência imediata desta admissão da hipótese da linguagem privada é que eu falo a minha linguagem privada L_p e Job fala a sua linguagem privada L'_p , eu atribuo o meu sentido em L_p à palavra «dor» e Job atribui o seu sentido em L'_p à palavra «dor», de modo que a minha única esperança de saber e de compreender o que Job quer dizer quando fala acerca de dor é transferir para o caso de Job aquilo que eu sei acerca do meu próprio caso quando utilizo a palavra «dor».

Esta transferência da minha própria experiência para a experiência de Job é justificada como sendo uma inferência *per analogiam* de tal modo que não me resta senão dizer que, como consequência do uso das nossas linguagens privadas, o meu conhecimento e a minha compreensão da experiência de dor de Job tem apenas um carácter analógico¹¹. Assim a tese a refutar tem a forma de uma proposição condicional, o antecedente da qual é a existência possível ou presumida de uma linguagem privada e o consequente da qual é a natureza analógica do meu conhecimento dos estados de consciência de outrem. O nosso propósito é agora expor os argumentos que Wittgenstein aduz contra

¹⁰ A mais conhecida tentativa de refutação do argumento de Wittgenstein aqui apresentado é a de Castañeda, Hector Neri «The Private Language Argument as a *reductio ad absurdum*» in *The Private Language Argument*, org. O. R. Jones, 1971, pp. 132 e segs.

¹¹ A tese de Wittgenstein representada no texto é o objecto do ensaio de Malcolm, Norman, «Knowledge of Other Minds» in *Knowledge and Certainty: Essays and Lectures*, Englewood Cliffs, New Jersey, 1963, pp. 130-140.

a natureza analógica deste conhecimento e desta maneira dar-nos a possibilidade de obter, por *modus tollens*, uma nova refutação da existência possível de uma linguagem privada como L_p .

Convém *in limine* regressar à formulação adoptada acima, nomeadamente de que sob a admissão da hipótese da existência de uma linguagem privada como L_p , o meu conhecimento e a minha compreensão dos estados de consciência de outrem é de natureza puramente analógica. Com efeito podemos aqui distinguir de facto duas proposições diferentes:

- i) A natureza do meu conhecimento dos estados de consciência de outrem é puramente analógica; e
- ii) É para mim destituído de sentido falar acerca dos estados de consciência de outrem.

É fácil de ver que a proposição ii é mais forte de que a proposição i no sentido que ii é condição suficiente de i uma vez que não posso sequer falar acerca da natureza do meu conhecimento dos estados de consciência de outrem, sem ter previamente uma ideia acerca do sentido das expressões que terão que ser empregues na descrição dos estados de consciência de outrem. Assim, a nossa intenção é agora proporcionar a análise da proposição ii e a partir dela estabelecer a conclusão desejada, nomeadamente de que não existe conhecimento analógico dos estados de consciências de outrem.

A primeira linha da argumentação que utilizamos é a do parágrafo 320 de *Philosophische Untersuchungen*, a qual consiste essencialmente em impedir a pretensão de analogicamente transferir para a experiência de dor de Job aquilo que foi ou é a minha própria experiência de dor. De facto se eu só posso chegar a compreender em que é que consiste a experiência de dor de Job a partir do modelo da minha própria experiência de dor, a questão surge como é que de facto a transferência se processa, uma vez que consiste em transferir a dor que de facto eu sinto para uma dor que de facto eu não posso sequer sentir. É relevante reparar que a transferência que se propõe que seja efectuada não é analógica no sentido de ser o género de transferência que normalmente eu consigo fazer na minha imaginação de transferir a sensação de dor, ocasionada por uma dor de ouvidos, para a sensação de dor, ocasionada por uma dor de dentes. No caso da transferência proposta eu não chego sequer a ter a dor que me proponho a transferir, uma vez que a dor é a dor de Job a que eu, *ex hypothesi*, não

tenho acesso directo. Não posso assim pois obter na minha imaginação as imagens mentais associadas com a dor de Job. Logo, se o sentido da palavra «dor» na minha linguagem L_p é estabelecido apenas por referência à minha percepção da minha própria dor, a própria palavra «dor» terá apenas o significado de designar a minha experiência privada de dor a que só eu posso ter acesso. Pela mesma razão ser-me-á também impossível poder falar acerca da experiência de dor de Job, uma vez que a palavra «dor» designa apenas na minha linguagem L_p a minha experiência privada de dor, ser-me-á impossível compreender o sentido da própria concepção de dor de Job.

Uma outra forma de argumentar a favor da concepção analógica acerca do meu conhecimento dos estados de consciência de outrem é recorrendo ao conceito, já utilizado anteriormente, de identidade. Durante a argumentação a favor da linguagem L_p apela-se, como foi então desenvolvido, para a identidade entre as sensações de dor que ocorrem em instantes, digamos I_1 e I_2 ,¹² a fim de justificar o emprego da mesma expressão em instantes diferentes. Uma variante desta concepção pode ser utilizada relativamente ao problema sob análise, nomeadamente apelando para a identidade das experiências de que eu e Job somos os sujeitos, no sentido em que eu me proponho agora a afirmar que a minha suposição de que Job tem uma experiência de dor é justificada pelo facto de eu com ela supor que Job tem o mesmo que eu tenho, quando tenho a experiência de dor.

Wittgenstein rejeita este uso do conceito de identidade para estabelecer a tese da natureza analógica do meu conhecimento dos estados de consciência de Job. É evidentemente claro que se eu tenho a experiência de dor e se Job tem a experiência de dor então é perfeitamente correcto afirmar-se que eu e Job temos a mesma coisa, nomeadamente a experiência de dor. Mas aqui supõe-se ter-se estabelecido independente e anteriormente cada um dos termos da conjunção, i.e., de que eu tenho a experiência de dor e que Job tem a experiência de dor e assim, a esta luz, a asserção da conjunção é completamente justificada. Todavia, no argumento agora proposto, conclui-se que eu e Job temos a mesma coisa sem ter havido previamente uma estipulação das condições debaixo das quais eu e Job seremos considerados como tendo a mesma experiência. Wittgenstein estabelece este ponto por intermédio do exemplo célebre da determinação da hora na Terra e no Sol, no

¹² Em outros pontos do texto referidos pela notação $t, k, t \cup k$.

parágrafo 350 de *Philosophische Untersuchungen*. Assim é perfeitamente claro o que eu digo quando afirmo que quando é na Terra meia-noite e no Sol meia-noite estou então justificado em afirmar que é a mesma hora na Terra e no Sol. Mas evidentemente na ausência de um critério¹³ para estabelecer as condições debaixo das quais se considerará que é a mesma hora na Terra e no Sol, a minha asserção que é no Sol a mesma hora que na Terra, i.e., meia-noite quando na Terra é meia-noite, permanece sem justificação e sem sentido.

Terminamos esta secção mostrando com mais detalhe como o recurso à noção de identidade não estabelece a pretensão da tese da natureza analógica do meu conhecimento dos estados de consciência de outrem. Trata-se agora de usar a sugestão do parágrafo 250 de *Philosophische Untersuchungen* e ver em que condições é que se pode propor estabelecer a identidade entre os predicados $D(x_i)$ ¹⁴ quando aplicados a mim e a Job. Como já foi indicado a questão revolve à volta do facto de a conjunção «eu tenho D e Job tem D » ter a sua asserção justificada quando já foi previamente decidido em que é que consiste a justificação de cada um dos conjuntos. O que não está justificado é o argumento que consiste em definir $D(x_i)$ em termos de um indivíduo, nomeadamente eu, e a seguir pretender estabelecer o sentido da expressão «Job tem dor» afirmando que se trata do mesmo predicado D que se usou na definição de «eu tenho D ». Logo, se $D(x_i)$ só pode ser definido em termos de um indivíduo, a sua aplicação a um outro indivíduo não pode ser feita sem se justificar a identidade do novo indivíduo e portanto o apelo para a identidade do predicado tem que ser considerado como irrelevante. A não ser assim, dever-me-ia ser então possível empregar, por exemplo, a expressão «a lua tem dor», e justificá-la dizendo que a lua tem o mesmo que eu tenho, nomeadamente dor. Portanto, se só me é possível estabelecer o sentido da palavra «dor» pela consideração da minha experiência privada de dor, a sua transferência para o caso da lua não pode ser feita apenas dizendo que a lua tem o mesmo que eu tenho. Logo a questão não é resolvida pelo simples recurso à identidade do predica-

¹³ O sentido do termo critério como o par $\langle K, P \rangle$ necessário para estabelecer a identidade entre os dois termos da conjunção é adequadamente definido na p. 87 e segs. de *Espontaneidade da Razão*.

¹⁴ A ideia geral de que ao ser humano x_i é atribuído o predicado «ter dor» é aqui representada na notação usual $D(x_i)$.

do. Não estamos assim de posse de um critério que nos permita estabelecer as condições debaixo das quais se considera que o mesmo predicado pode ser aplicado, e à falta deste critério se reduz a falta de sentido da pretensão da tese analógica do meu conhecimento acerca dos estados de consciência de outrem.

4. A refutação da tese da inserção subsequente de L_p em L

Dois novos aspectos do nosso problema acerca da natureza do conhecimento dos estados de consciência de outrem constituirão o objecto das considerações a desenvolver neste parágrafo, nomeadamente a questão que se refere à determinação do lugar da experiência interior como fundamento da lógica das expressões usadas acerca de sensações, como a sensação de dor e ainda a determinação do lugar da experiência interior como fundamento do grau de conhecimento que é possível alcançar acerca dos estados de consciência de outrem. Em particular, a nossa primeira questão pode ser mais dramaticamente expressa pela determinação do grau de prioridade da linguagem L_p em relação a uma linguagem intersubjectiva a que chamaremos L . Partindo da premissa da existência assumida de L_p parece ser plausível argumentar que o sentido de uma expressão de L é função do sentido da mesma expressão em L_p e que é portanto a inserção subsequente de L_p em L que constitui L .

Com efeito, segundo a concepção da inserção subsequente de L_p em L , a palavra «dor» adquire para mim o seu sentido na medida em que eu a concebo como o nome geral de uma experiência, à concepção da qual eu só chego através da operação de abstracção¹⁵ efectuada sobre minhas próprias experiências de dor, e por isso é que só em sentido puramente analógico eu consigo chegar a compreender em que é que consiste a experiência de dor de Job. A nossa questão consiste agora pois em estabelecer se esta operação nos proporciona o processo que precisamos para compreendermos o sentido da palavra «dor» em L , i.e., na linguagem suposta ser intersubjectiva. Por um lado, como se pode deduzir dos resultados do capítulo anterior, a palavra «dor» é apenas uma das partes constituintes do jogo de linguagem acerca de dor e da sua expressão, de modo que o simples uso

¹⁵ No sentido acima referido em que o termo «dor» é entendido como o predicado $D(x)$.

desta palavra em si não é suficiente para estabelecer se estou ou não de posse do jogo de linguagem em questão. Mas além desta uma questão mais fundamental é agora posta à nossa consideração, nomeadamente que a justificação da inserção de L_p em L não foi ainda oferecida de modo a ser esclarecido o desejado processo de transferência, a efectuar na minha consciência, do sentido da palavra «dor», sugerido pela minha própria experiência, para a experiência de outrem. Esta transferência parece *prima facie* relativamente simples de conceber e assim de ser justificada, uma vez que eu e Job usamos a palavra «dor» e para cada um de nós parece que a palavra adquire o seu sentido por estar imediatamente associada com a experiência que eu (respectivamente nós) tenho (respectivamente temos) de dor, de tal modo que a ela devem estar associadas as diversas imagens mentais que subsequentemente representam a experiência de dor.

Aceite esta interpretação não se pode no entanto deixar de ser levado a encarar a questão da possibilidade de Job ter da experiência de dor uma experiência qualitativamente diferente da minha, e portanto a dúvida acerca da relação da palavra «dor» e a experiência de dor torna-se legítima, no sentido em que eu não pareço ter meios de saber se falamos acerca da mesma experiência quando ambos usamos a palavra «dor». Como não me é possível entrar na consciência de Job a fim de poder determinar com completa segurança que género de experiência é que Job tem, quando tem a experiência a que ele chama «dor», não me resta, nesta introspecção, senão estipular que todos os utentes da linguagem chamem precisamente sempre «dor» à mesma sensação. No entanto, à luz da interpretação proposta, eu só posso chegar a ter conhecimento do sentido da palavra «dor» através da consideração da minha própria experiência de dor, e é portanto agora possível que aquilo a que eu acabo de chamar «dor» seja qualitativamente diferente daquilo que Job chama pelo mesmo nome.

Não é agora possível calar esta dúvida, i.e., a dúvida acerca da identidade do que eu e Job temos quando ambos usamos a palavra «dor», dizendo que seja como for temos ambos pelo menos qualquer coisa a que estamos dispostos a chamar «dor», uma vez que, *ex hypothesi*, esta qualquer coisa que ambos temos é presumivelmente também privada, no sentido em que foi estabelecido o emprego deste termo nos pará-

grafos anteriores, e assim volta a ser possível que o que ambos temos seja diferente¹⁶.

Concluimos assim que a nossa hipótese de que o sentido de uma expressão de L depende do sentido da expressão em L_p , após a inserção subsequente de L_p em L , representa um desvio das estipulações apresentadas no Capítulo 1 de *Espontaneidade da Razão* acerca da aquisição da competência linguística através de jogos de linguagem, no nosso caso num jogo de linguagem acerca do emprego da palavra «dor». De facto, como vimos então, é devido à circunstância de as palavras que se referem às experiências da consciência pertencerem *ab initio* a L , que se consegue falar acerca destas experiências em L . Logo a inserção subsequente de L_p em L não pode ser o que constitui o sentido de uma expressão em L .

Wittgenstein cristalizou estas considerações na conhecida imagem do parágrafo 293 de *Philosophische Untersuchungen*. Aqui supõe-se, para efectuar a comparação, que eu e Job temos cada um uma caixa dentro da qual está algo a que ambos chamamos «escaravelho». Eu não estou em condições de ver a caixa de Job e Job não está em condições de poder ver a minha caixa. Assim agora cada um de nós sente-se justificado em afirmar que o sentido da palavra «escaravelho» é estabelecido para cada um de nós apenas pela observação do seu próprio escaravelho. Mas neste caso, uma vez que o sentido da palavra «escaravelho» é estabelecido apenas pela consideração do meu próprio «escaravelho», eu posso chamar «escaravelho» ao que quiser; em particular a minha caixa pode conter um polvo em vez de um escaravelho, ou o que eu tenho na caixa é um objecto em transformação, ou simplesmente a caixa está vazia no momento. Se na nossa linguagem L a palavra «escaravelho» tem um emprego determinado, não pode este emprego ser concebido como sendo o nome do objecto dentro da minha caixa. Este objecto é irrelevante para o género de afirmações que eu me encontro a fazer e portanto, segundo Wittgenstein, não está incluído neste jogo de linguagem — a caixa até pode estar, como já foi mencionado, vazia. Como é óbvio a caixa representa a consciência e o escaravelho a sensação de dor, tal como são representadas na doutrina na concepção analógica do meu conhecimento acerca dos estados de consciência de outrem. A imagem (do escaravelho na caixa ou) da

¹⁶ Cf. Malcolm, Norman, «Wittgenstein's Philosophical Investigations» in *Knowledge and Certainty: Essays and Lectures*, Englewood Cliffs, New Jersey, 1963, pp. 96-129.

natureza privada dos objectos da minha consciência recebe o seu poder de atracção de principalmente duas fontes: a já mencionada falácia da analogia e a doutrina, também já mencionada, acerca da natureza da linguagem segundo a qual a função essencial desta é a função de nomear.

Quanto à primeira resta-nos observar que aqui a falácia da analogia consiste em tentar transferir para o universo da consciência as formas de representação usadas no universo sensível, nomeadamente a análise deste em termos de objectos e de processos. Assim deve-me ser possível encontrar na minha consciência os objectos e processos que possibilitarão a sua análise e que proporcionarão a desejada referência para as expressões *L*, as quais podem ser interpretadas como asserções acerca da consciência. Quanto à segunda destas fontes, apelamos uma vez mais para os resultados do Capítulo 1 de *Espontaneidade da Razão* e para a crítica de Wittgenstein ao que ele chamou de alimentar o pensamento à custa de uma dieta unilateral. Não nos resta pois senão recordar que, se se interpreta a função da linguagem à luz de uma função única de designar objectos do chamado mundo exterior, então torna-se irrelevante procurar distinguir entre as diversas categorias de objectos assim designados, obtendo-se assim a mesma interpretação do funcionamento da linguagem considerando objectos em si tão diferentes como dores, números, cores ou direcções. Assim se interpretamos a natureza lógica das expressões que utilizamos para falar acerca de sensações à luz do modelo único de designação e objecto, a consideração do objecto neste género de expressões torna-se irrelevante.

Voltamo-nos agora para a segunda das nossas questões, nomeadamente a questão acerca do lugar da experiência interior na determinação do grau de conhecimento que eu posso ter acerca da consciência de outrem, em particular no que se refere à inacessibilidade dos estados de consciência de outrem. Wittgenstein dramatizou com sugestivo realismo a natureza desta inacessibilidade nas suas diversas discussões acerca do princípio do terceiro excluído. Assim para eu poder pretender afirmar com sentido que «ou existe um número inteiro com a propriedade *P* ou tal número não existe» eu tenho que supor, para aceitar a verdade da minha declaração, a existência de uma consciência essencialmente diferente da minha, nomeadamente uma consciência infinita, que tem a percepção simultânea da totalidade dos números inteiros, e que à luz desta percepção pode decidir qual é o número inteiro com a propriedade desejada, ou então que não existe na totalidade dos números inteiros um número com a

propriedade desejada. Uma pretensão idêntica acerca dos poderes da minha percepção parece também encontrar-se na minha declaração «ou Job tem a sensação de dor ou Job não tem a sensação de dor» onde de novo eu tenho que supor uma extensão dos poderes da minha percepção, a qual não me é normalmente dada na experiência, se quero afirmar que a minha disjunção tem um sentido determinado. Mas em ambos os casos, como Wittgenstein indica no parágrafo 352 de *Philosophische Untersuchungen*, embora o sentido das minhas declarações pareça estar estabelecido inequivocamente, esta determinação de sentido depende essencialmente de metáforas acerca da acessibilidade dos objectos do pensamento, no primeiro caso na percepção simultânea da infinitude da totalidade dos números inteiros e, no segundo, na percepção dos objectos da consciência de Job. Em vista destas circunstâncias parece então razoável argumentar a favor da natureza analógica do meu conhecimento dos estados de consciência de outrem, nomeadamente usando o facto do aparente contraste entre o grau de conhecimento alcançado: enquanto que eu posso saber de mim próprio se tenho ou não uma experiência de dor, acerca de Job não posso mais do que supor que ele tem a experiência de dor.

Contra esta forma de argumentação começa Wittgenstein por observar que, compreendendo o termo «saber» tal como é normalmente empregue, é um facto da experiência que eu sei se ou quando Job está a ter uma experiência de dor, e portanto as alegadas dificuldades e graduações epistemológicas acerca da natureza deste conhecimento só podem ser devidas a um desvio do sentido do termo tal como é normalmente empregue. Por isso a relacionada pretensão de que Job não pode saber acerca de mim se eu tenho a sensação de dor com a certeza de que só eu posso ter, é igualmente rejeitada à luz da interpretação anterior do termo «saber». A esta luz reconhece-se que não faz sentido tentar dizer de mim que eu sei que tenho a sensação de dor uma vez que, neste sentido, o meu saber que tenho a sensação de dor se reduz ao simples facto de eu ter a sensação de dor.

De facto a tese da natureza analógica do meu conhecimento dos estados de consciência de outrem consiste, como vimos, em manter que a palavra «dor» é o nome de uma determinada experiência interior, nomeadamente a ocorrência da sensação de dor na minha consciência, à qual ninguém, a não ser eu próprio, pode ter qualquer acesso. Assim quando Job diz de si próprio que tem a experiência de dor, é esta experiência, que eu conheço do meu próprio caso, que é o sentido desta declaração. Quando Job, por sua vez, diz de mim que

eu tenho a experiência de dor, quer com isso dizer que eu tenho na minha consciência a ocorrência do mesmo género de sensação que ele tem quando fala das suas experiências de dor. Logo, no meu caso, o meu conhecimento da minha sensação de dor é um dado imediato da consciência e, como não posso ter acesso à consciência de Job, não posso alcançar mais, acerca do que se passa na sua consciência, do que uma conjectura.

O que Wittgenstein rejeita nesta resposta é que se a suposição fosse aceite de que só acerca de mim próprio é que posso saber se tenho ou não a sensação de dor e que acerca de Job não posso fazer mais do que uma conjectura, então o emprego do termo «saber» sofreu uma modificação que não foi legitimada na argumentação proposta no argumento por analogia, uma vez que, no sentido aceite, é normalmente um facto da experiência que eu sei com relativa segurança quando Job tem uma sensação de dor e quão destituído de sentido se me apresenta a noção de eu saber acerca de mim próprio que tenho a sensação de dor.

5. A eliminabilidade de expressões de autoconhecimento

As considerações que até aqui representámos acerca da natureza da experiência interior não são por si suficientes para conduzirem a uma decisão acerca de um problema tradicional da filosofia da consciência, nomeadamente o problema do dualismo psicossomático, i.e., a partição do mundo da experiência em consciência e realidade, ou corpo e espírito ou sujeito e objecto. Embora Wittgenstein não aborde estas questões na sua formulação tradicional, a sua doutrina acerca destas questões pode no entanto ser deduzida das suas considerações acerca da estrutura geral que estas questões apresentam em comum. Trata-se nelas, de facto, de determinar as relações recíprocas entre o mundo interior e o mundo da experiência fenomenal o que, do ponto de vista de Wittgenstein, consiste em determinar o sentido das expressões que se referem, na linguagem corrente, às experiências interiores. A cada uma destas expressões que, na linguagem corrente, representa uma destas experiências, chama Wittgenstein uma exteriorização¹⁷.

¹⁷ Acerca da relação entre este problema e o problema tradicional ver Hacker. P. M. S. «A Cloud of Philosophy Condensed into a Drop of Grammar» in *Insight and Illusion*, Clarendon Press, Oxford, 1972, pp. 252 e segs.

Assim, uma exteriorização é uma declaração feita por mim de que uma dada experiência interior está a decorrer, um exemplo típico de uma exteriorização é minha declaração «dói-me a cabeça» ou qualquer das suas variantes, como «tenho uma dor de cabeça». O problema com que nos defrontamos agora é o de tentar elucidar a natureza lógica destas expressões; em particular queremos nesta secção expor a doutrina de Wittgenstein acerca da natureza lógica das exteriorizações sob uma forma que conduza a uma decisão acerca do problema tradicional mencionado acima. Finalmente teremos que determinar a posição relativa desta doutrina em relação ao argumento já apresentado contra a existência de uma linguagem privada.

A doutrina de Wittgenstein deixa-se formular sucintamente na seguinte proposição: uma exteriorização não é uma proposição e, assim, o princípio do terceiro excluído não se lhe aplica. Mas o que não pode ser verdadeiro nem falso também não pode ser objecto de conhecimento. Logo que é representado na exteriorização é incognoscível. Por isso não tem sentido pretender afirmar «eu sei que tenho uma dor de cabeça». Deste modo pretende Wittgenstein reduzir o conhecimento que eu afirmo que tenho acerca de minha dor à própria dor. Chegamos assim à conclusão que o conhecimento que eu tenho acerca dos meus estados de consciência não é uma parte essencial da minha explicação do sentido das expressões em que eu afirmo conhecer os meus estados de consciência, e por isso eliminável da minha explicação do sentido dessas expressões.

Para expor esta nova tese, a que chamaremos a tese acerca da eliminabilidade da expressão de autoconhecimento¹⁸, começaremos por regressar ao ponto já alcançado anteriormente, de que rejeitada a existência da linguagem L_p só faz sentido falar acerca da linguagem intersubjectiva L , em relação à qual consideramos como estabelecido *in limine* o segmento de L que se refere às sensações. Este segmento, à luz do argumento aduzido contra a existência de L_p não é estabelecido como uma parte de L cuja função é descrever as ocorrências de sensações na minha consciência mas antes, ao contrário, a concepção destas sensações só me é acessível através da aquisição da linguagem L . Por outro lado, como podemos deduzir de um dos nossos resultados da Teoria do Sentido, não podemos esperar que o termo «dor» tenha o

¹⁸ Prefixos como «eu sei que» podem ser designados por operadores epistémicos, e a tese de Wittgenstein é pois que operadores epistémicos, aplicados a descrições de estados de consciência, se reduzem aos próprios estados de consciência.

mesmo sentido, e portanto a mesma forma de aplicação, em sentido absoluto. Como foi então proposto, este termo terá também que ser contextualmente relativizado. Wittgenstein distingue dois contextos que determinam duas formas essencialmente diferentes da aplicação, e portanto do sentido, de um termo como «dor», nomeadamente o contexto em que o termo ocorre como predicado de uma proposição cujo sujeito é a primeira pessoa do presente do indicativo, e um outro contexto em que o termo ocorre como predicado numa proposição cujo sujeito é a terceira pessoa do presente do indicativo.

A tese acerca da eliminabilidade da expressão acerca do autoconhecimento depende de uma premissa geral que é agora útil formular: aquilo de que eu não posso duvidar também não posso ter a certeza e aquilo de que não posso ter a certeza também não posso conhecer. À luz desta premissa podemos agora propor a generalização de uma ideia já anteriormente mencionada acerca da eliminabilidade do prefixo «eu sei que» quando é aposto a uma exteriorização. Considerando um exemplo já mencionado, «eu tenho uma dor de cabeça» o resultado de prefixar «eu sei que» é a aparente proposição «eu sei que eu tenho uma dor de cabeça». Mas agora a expressão do meu conhecimento, tal como a expressão da minha dúvida, torna-se destituída de sentido, uma vez que aquilo de que não faz sentido eu duvidar também não faz sentido eu pretender ter conhecimento. Não é esse obviamente o caso quando o termo «eu» em «eu tenho uma dor de cabeça» é substituído por «ele». Eu posso perfeitamente afirmar «eu sei que ele tem uma dor de cabeça» como também posso perfeitamente afirmar «eu duvido que ele tenha uma dor de cabeça»; analogamente, o desejo de conhecer é tratado nesta análise do mesmo modo, e assim faz sentido afirmar «eu gostaria de saber se ele tem uma dor de cabeça» e não faz sentido afirmar «eu gostaria de saber se eu tenho uma dor de cabeça». O mesmo género de análise pode ainda ser aplicado a designações temporais, de tal modo que faz sentido eu dizer «eu sei que ontem tive uma dor de cabeça» e já não faz sentido eu dizer «eu sei que agora tenho uma dor de cabeça».

Embora seja possível dividir as exteriorizações que são expressas na primeira pessoa do presente quanto à sua função em duas classes principais, aquelas cuja função é comunicar descritivamente o conteúdo de uma experiência interior e aquelas cuja função é exprimir essa experiência, a análise de Wittgenstein é essencialmente concebida para tratar das exteriorizações desta segunda classe, e é para estas, pois, que é proposta a tese da eliminabilidade da expressão do autoconhecimento.

Para compreender a natureza puramente expressiva das exteriorizações, convém recordar os resultados da sinopse da Teoria do Sentido, e recordar que o sentido de um termo como «dor» é compreendido quando se compreendem as circunstâncias e o contexto com que a introdução do termo em *L* está associada. A ideia de Wittgenstein é que a introdução do termo «dor» traz como consequência a introdução de uma nova forma de comportamento a ser observável durante a experiência de dor, em particular, o termo «dor» é introduzido de facto como uma substituição das expressões orgânicas e inarticuladas de dor, como gritar ou chorar. Assim, segundo o parágrafo 224 de *Philosophische Untersuchungen* o termo da linguagem *L* não é uma descrição do comportamento ou da experiência de dor, mas é antes para ser considerado como tendo a natureza lógica de um grito, que o termo de resto é suposto substituir e, portanto não pode ser interpretado como uma asserção à qual um valor de verdade será atribuído. Assim, à luz da nossa premissa, não faz sentido duvidar se Job está ou não enganado quando diz que está em dor ou, o que seria o mesmo, não faz sentido perguntar se Job não se terá enganado ao chorar a sua dor¹⁹.

Presumida a existência de um comportamento associado com a experiência de dor, em si próprio orgânico, seria de concluir a impossibilidade de ensinar ou introduzir em *L* um termo como «dor», cujo sentido é essencialmente determinado pelo facto de podermos associar o termo com as diferentes manifestações orgânicas de dor, e em relação às quais justificamos a introdução e a explicação do sentido do termo. Como é tornado óbvio no parágrafo 288 de *Philosophische Untersuchungen*, a explicação do sentido do termo não é por si suficiente para garantir a assimilação correcta do termo, uma vez que qualquer explicação pode ser correcta ou incorrectamente assimilada. Só o comportamento subsequente da pessoa a quem damos esta explicação do sentido do termo «dor» pode levar a decidir se a sua assimilação foi correcta ou incorrectamente realizada. Finalmente recordamos que não podemos regressar à concepção já rejeitada de que a compreensão do sentido deste termo é explicável por referência ao acto de visualização que ocorre na minha consciência quando eu emprego o termo, uma vez que, como sabemos, nem a experiência

¹⁹ A relação entre o conhecimento de si próprio e o emprego do conceito de critério entendido como o par $\langle K, P \rangle$ é elegantemente exposto por Shoemaker, Sydney, «Self-Reference and Self-Awareness» in *The Journal of Philosophy*, volume LXV, n.º 19, 1968.

privada de dor nem a imagem mental privada são essenciais na elucidação do sentido do termo. No parágrafo 271 de *Philosophische Untersuchungen* refere-se Wittgenstein, num exemplo, a uma pessoa que não conseguisse conservar na sua memória o sentido do termo «dor», que utilizasse o termo «dor» a propósito dos objectos mais variados, mas que no entanto também o empregasse em conformidade com o seu uso e as pressuposições correntes²⁰. No exemplo, o sentido que a pessoa não consegue conservar na memória não é pois o sentido do termo como o delineamos no nosso primeiro capítulo, mas o sentido tal como é incorporado numa imagem mental privada que a pessoa tem que recordar para poder aplicar o termo. Mas uma vez que a utilização do sentido do termo em situações reais pode, *ex hypothesi*, variar de objecto para objecto sem que no entanto a pessoa deixe de observar o seu emprego correcto, deixa de ser necessário pressupor a hipótese redundante de que o sentido do termo não é conservado na memória e, assim, a sua compreensão não pode também depender desta circunstância.

Wittgenstein considera um outro aspecto da análise de declarações da forma geral «*X* tem dores», nomeadamente o estatuto lógico do sujeito da função proposicional «*X* tem dores». Considerando que *X* toma valores num domínio de seres humanos, Wittgenstein procura determinar as consequências que se seguem quando *X* é substituído por «eu» ou por «Job»²¹. Em particular quando *X* é substituído por «eu» e eu afirmo, «eu tenho dores» não é necessariamente o caso que eu saiba quem tem as dores que eu digo ter. Em conformidade com o nosso postulado de que aquilo de que não posso ter dúvidas também não posso ter a certeza e aquilo de que não posso ter a certeza também não posso conhecer, segue-se que na expressão «eu tenho dores» o termo «eu» não designa uma pessoa, uma vez que não existe um critério que identifique a pessoa de que eu afirmo ser o sujeito da experiência das dores, visto que, de mim próprio, não posso duvidar se estou ou não a ter as dores que digo que tenho.

²⁰ Uma variante do problema da inversão do espectro, em que se supõe que o sentido privado das palavras que têm como referência as cores do espectro difere de pessoa para pessoa, embora se consiga alcançar acordo quanto ao uso intersubjectivo das palavras em questão. Logo, o uso corrente é conservado mas as imagens mentais privadas diferem uniformemente.

²¹ A discussão mais completa do tema na literatura presente é Shoemaker, Sydney, *Self-Knowledge and Self-Identity*, Cornell University Press, Ithaca, Nova Iorque, 1963.

Em contraste com as exteriorizações, as declarações expressas na terceira pessoa do singular são autenticamente descritivas e não apenas expressivas, e por isso tem sempre sentido perguntar-se de quem é que se fala quando se diz que Job tem dores e o que é que se lhe atribui ao dizer-se que Job tem dores. Assim a proposição «Job tem dores» não pode pois ser considerada como uma reformulação da expressão espontânea e orgânica de dor de Job e por isso tem que se determinar a referência da palavra «dor» neste contexto. Excluída a possibilidade de explicar a referência da palavra «dor» apelando para a experiência interior de Job, à qual só ele pode ter acesso, não resta senão tentar pô-la em conexão com o comportamento de dor que Job ostensivamente mostra. No entanto, este comportamento, como já tivemos ocasião de fazer notar, não é para ser interpretado à luz da doutrina behaviorista, mas antes como sendo a união do comportamento observável e do uso de um critério expresso na linguagem L que permite, por um lado, a identificação do sujeito da experiência de dor como sendo Job e, por outro lado, permite a identificação de «dor» como sendo o conteúdo da experiência de que Job é sujeito. Nestas circunstâncias, pois, a declaração feita com o emprego da terceira pessoa do singular é para ser considerada como uma proposição genuína e portanto susceptível de ser portadora de um valor de verdade. Neste caso também o nosso postulado conduz a uma decisão acerca da questão de determinar o sentido do meu conhecimento da dor de Job, uma vez que, acerca de Job, já faz sentido eu duvidar quem é o sujeito da experiência e/ou de que experiência se trata. Esta dúvida uma vez eliminada conduz à declaração verificavelmente verdadeira de que Job tem a experiência de dor.

Concluimos este parágrafo pela alusão ao problema da ordem entre o argumento contra a existência da linguagem L_p , e a doutrina da eliminabilidade da expressão de autoconhecimento. O problema em si consiste em saber se as proposições em questão são deduzíveis uma da outra e, em caso afirmativo, qual é a proposição que se utiliza como premissa. Este problema não seria sequer posto se, como facto da experiência, as proposições que exprimem uma pretensão de se ter alcançado autoconhecimento tivessem um único emprego, nomeadamente, o de funcionarem como uma expressão de dor. Ao contrário, é um facto da experiência linguística que eu posso usar uma dessas proposições, como «eu sei que tenho uma dor de cabeça» num contexto em que se torna óbvio que não estou a ter a pretensão de ter alcançado conhecimento acerca de mim próprio. A proposição pode

ser apenas usada, por exemplo, para exprimir um motivo ou um fundamento justificativo da minha actuação, como quando digo «eu sei que tenho uma dor de cabeça» para justificar o motivo pelo qual não devo aceitar a proposta de ir sair.

Tem-se igualmente como um facto da experiência linguística que o emprego deste género de proposições poder ter um fim genuinamente descritivo, como é o caso quando eu tenho que relatar ao meu médico a configuração precisa da minha dor de cabeça a fim de poder vir a receber um tratamento adequado. Para estas proposições temos como um resultado geral que a doutrina da eliminabilidade da expressão do autoconhecimento é falsa, uma vez que para estes casos temos que considerar as proposições como sujeitas ao princípio do terceiro excluído. Mas do facto de a doutrina ser falsa para estas proposições não se infere que o argumento contra a existência de L_p seja falso, uma vez que o que se está a garantir a estas proposições é a representabilidade em L . Ao aceitarmos a representabilidade destas proposições em L não estamos, de modo algum, a afirmar a existência possível de L_p , dado que é possível, à maneira de Kant, manter que uma condição de possibilidade da minha experiência é o reconhecimento da minha possibilidade de imputar experiências a outrem. Deste modo posso definir um contraste radical entre o uso de L e o uso de L_p , uma vez que, relativamente a L , eu sei que tenho determinada experiência e que Job tem uma determinada experiência, e o meu saber acerca de mim é o mesmo que o meu saber acerca de Job. Por contraste, em L , só posso saber acerca de mim que tenho uma determinada experiência e tenho que ser céptico relativamente à minha possibilidade de o saber acerca de Job. Logo só posso negar a possibilidade de conhecimento acerca de Job se nego a possibilidade de conhecimento acerca de mim próprio. Assim para determinar o que se passa na consciência de Job, usando uma inferência tradicional *per analogiam* só me é possível quando eu estou em condições de determinar o que se passa na minha própria consciência. Mas evidentemente se me é possível saber o que se passa na minha própria consciência não preciso da inferência *per analogiam* para estabelecer o que se passa na consciência de Job.

Finalmente para aquelas proposições que exprimem uma pretensão actual de autoconhecimento, a validade do argumento contra a existência de L_p implica a verdade da doutrina da eliminabilidade da expressão de autoconhecimento. Com efeito, como vimos durante a exposição do argumento contra a existência de L_p , este argumento depende da rejeição da pretensão da privacidade epistemológica. Se a

doutrina acerca da eliminabilidade da expressão de autoconhecimento fosse falsa, então seria possível interpretar estas proposições como sendo expressões de conhecimento a que só o sujeito poderia em princípio ter acesso. Estas proposições seriam então inevitavelmente apenas representáveis em L_p . Mas em L_p , *ex hypothesi*, não tem existência possível. Logo se não pode haver uma linguagem como L_p também não pode haver privacidade epistemológica, e não havendo esta também não me é dado saber se tenho ou não a experiência que digo estar a ter. Portanto a doutrina da eliminabilidade da expressão de autoconhecimento é verdadeira para as proposições que exprimem uma pretensão actual de autoconhecimento se o argumento contra a existência de L_p , for verdadeiro.

Em relação à primeira classe de proposições, i.e., aquelas em que se pode inferir do contexto que o princípio do terceiro excluído se lhes aplica, a doutrina da eliminabilidade da expressão do autoconhecimento não é aplicável e o problema de sua derivação a partir do argumento contra a existência de L_p não se coloca.

6. A relação de critério $\langle K_i, p \rangle$ e o princípio $D_i(S) \Leftrightarrow D_{i \cup k}(S)$

Concluimos com uma referência ao emprego do termo «critério» feito por Wittgenstein ao longo do argumento contra a existência da linguagem L_p e da sua doutrina acerca da eliminabilidade da expressão de autoconhecimento.

O aspecto mais imediatamente óbvio em relação ao emprego do termo «critério» por Wittgenstein é a introdução do contraste entre um critério e um sintoma. Num exemplo conhecido, a baixa de pressão num barómetro pode ser usada como um sintoma de chuva; analogamente, as transformações eléctricas e químicas das células do meu cérebro podem ser usadas como um sintoma de que estou a ver um cravo. A baixa pressão do barómetro e as transformações eléctricas e químicas do meu cérebro não são, no entanto, critérios que permitem determinar se chove ou se estou a ver um cravo. A chuva torna-se manifesta para mim através das minhas diversas impressões e adaptações que me são forçadas pelas circunstâncias; analogamente, apercebo-me de que tenho diante de mim um cravo sem necessariamente ter que ter uma ideia da configuração eléctrica e química do meu cérebro neste momento. Assim um critério é uma norma de

procedimento, incluindo o uso da linguagem, por meio da qual se torna possível determinar com decisiva evidência se estou ou não em condições de poder afirmar uma proposição dada P^{22} . A relação entre a proposição P e um critério C_i para P não é unívoca uma vez que, à luz do parágrafo 164 de *Philosophische Untersuchungen*, os critérios dependem das circunstâncias e em condições diferentes definem-se critérios diferentes.

Como já tivemos a ocasião de ver, um critério para identificar uma sensação é uma norma de procedimento que nos leva a poder dizer de que sensação se trata; ou, num outro exemplo, um critério para declarar duas funções como idênticas é uma norma de procedimento que nos leva a poder dizer que as duas funções são as mesmas; um critério para decidir da correcção de uma memória é uma norma de procedimento que nos leva a poder dizer se a memória é ou não a memória correcta.

Um critério difere de um sintoma pelo facto de ser uma norma decisiva de evidência. Num exemplo conhecido de Wittgenstein em *Philosophische Bemerkungen*, um critério para se poder determinar se eu sofro ou não de angina é a descoberta do bacilo no meu sangue, de tal modo que dizer «eu tenho angina se tiver bacilo» é uma tautologia; assim, ao encontrar-se o bacilo encontrou-se a evidência decisiva para se poder justificar a minha afirmação «tenho uma angina». Mas a noção de critério como evidência decisiva tem que ser reconciliada com a noção anteriormente mencionada de que a relação criterial não possui a propriedade de univocidade; assim, se um critério é evidência decisiva para poder afirmar uma dada proposição P , parece seguir-se que não pode haver mais do que um critério. A demonstração desta aparente impossibilidade é fácil conduzir; basta apelar nomeadamente para o facto de que a existência de, digamos, dois critérios, sugere imediatamente a questão de saber se são ou não entre si inconsistentes, uma hipótese para qual não parece haver uma razão para ser rejeitada *in limine*. Se os critérios estão em conflito então um pelo menos não pode ser considerado como proporcionando evidência decisiva, o que demonstra a nossa proposição. Vale a pena reparar que o mesmo argumento obtém quando se substitui a expressão «evidência decisiva» por a «melhor evidência possível».

²² Acerca do uso da relação de critério na refutação da existência da linguagem L_p ver Kenny, Anthony, «Criterion» in *Encyclopaedia of Philosophy*, org. Paul Edwards, Nova Iorque, pp. 258 e segs.

O processo que é de presumir que Wittgenstein tem em mente é o de conceber a relação criterial, i.e., a relação entre os dois critérios putativos K_1 e K_2 e uma proposição P que eu me considero à sua luz justificado em afirmar, como sendo a relação que se obtém quando de K_1 ou K_2 se pode afirmar P e de não- K_1 e não- K_2 não se segue que eu afirme não- P ; logo, é uma relação segundo a qual K_1 e K_2 não estabelece necessariamente P e, por sua vez, não- K_1 e não- K_2 não estabelece necessariamente não- P ²³. Por outro lado, através de exemplos dados por Wittgenstein em *Philosophische Untersuchungen*, chega-se à noção de que aquilo que eu digo é efectivamente um critério para determinar as imagens mentais que eu tenho neste momento na minha consciência; mas o facto de eu dizer que neste momento tenho uma imagem mental do Sol na minha consciência não estabelece decisivamente que eu tenho uma imagem mental do Sol neste momento na minha consciência, uma vez que eu posso estar enganado ou a mentir deliberadamente. Como vimos durante o argumento contra a existência de L_p , uma memória não pode ser um critério para a identificação de uma sensação visto que uma memória não é evidência decisiva para estabelecer a existência da sensação, uma vez que cada memória requer por sua vez um critério para estabelecer sua correcção. Assim concluímos que a relação criterial não é a relação de implicação.

A expressão «norma de procedimento» pode agora ser considerada com mais pormenor. Com ela temos em mente a ideia de que um critério é para ser agora concebido como um critério definidor e este por sua vez consiste num método de determinação do fenómeno que fornece a evidência decisiva que procuramos. No parágrafo 354 de *Philosophische Untersuchungen* Wittgenstein afirma que a existência simultânea de critérios e de sintomas arrasta consigo a consequência indesejável de obliterar o conceito de critério, criando a impressão falsa de que só há sintomas. Num exemplo já mencionado, a experiência mostra que quando a pressão baixa há chuva, mas também mostra que há chuva quando eu me sinto molhado. Normalmente rejeita-se a minha impressão de que estar molhado é um critério de ter chovido pelo facto de os dados dos sentidos poderem ser uma fonte de erro. Mas como é que eu me posso dar a ilusão de que chove senão pelo uso do critério definidor? É óbvio que os dados imediatos

²³ Nestes termos a relação de critério poderia ser representada pela conjunção $((K_1 \vee K_2) \rightarrow P) \wedge \sim((\sim K_1 \wedge \sim K_2) \rightarrow \sim P)$.

dos sentidos não são evidência decisiva de chuva. Mas a diferença entre critérios e sintomas consiste em que o valor evidencial dos sintomas é apreendido na experiência enquanto o valor da relação criterial enquanto evidência decisiva depende de uma definição. Se K_i é um critério para P então é necessário que K_i é evidência para P . Como a notação implica, pode haver mais do que um critério para P . Logo K_i é um critério para P se e somente se K_i constitui evidência imediata e não indutiva da verdade de P .

Resta observar finalmente, para caracterizar o uso do termo, que ao falar-se de critério se tem em mente não só o fenómeno ou fenómenos que são para ser considerados como proporcionando a evidência decisiva que procuramos, mas também a representação verbal do critério definidor, o que implica um uso do termo que é idêntico ao uso do termo «critério» em matemática. Assim a proposição necessariamente verdadeira que representa a relação de evidência entre K_i e P é também para ser considerada como sendo um critério para P . A natureza da relação que Wittgenstein supõe existir entre K_i e P é puramente convencional. K_i é estabelecido por mim num acto livre de escolha e pode ser ab-rogado por mim quando eu quiser. Mas K_i não é um critério para P se eu conseguir chegar a compreender o sentido de P sem ter compreendido previamente a relação que existe entre K_i e P .

O conceito de critério, na filosofia da consciência de *Philosophische Untersuchungen*, é normalmente identificado com um fenómeno observável, o que necessariamente é evidência da existência de um estado de consciência, ou de um processo de consciência em si próprio inobservável. Deste modo, o critério para determinar a cor da imagem que Job tem em mente neste momento é o que eu oiço e vejo Job fazer, uma vez que não posso observar directamente a cor da imagem que Job diz ter. Em geral, segundo o parágrafo 269 de *Philosophische Untersuchungen*, temos ao nosso dispor critérios que permitem deduzir, do comportamento de Job, o facto de ele pensar que compreende o sentido da expressão que eu uso, o facto de ele atribuir à minha expressão um sentido que não é o sentido correcto e finalmente o facto de compreender correctamente o sentido da expressão que eu uso. A ideia geral é que a determinação da existência de um processo da consciência só pode ser feita recorrendo a um critério no sentido do parágrafo 269 de *Philosophische Untersuchungen*. Vale a pena reparar que embora os critérios para a aplicação de um termo sejam formulados em termos de comportamento, não se segue necessariamente que o termo assim introduzido seja ele próprio uma descrição

de comportamento. Como já tivemos a ocasião de expor, o facto de K_i ser um critério para P não é para ser identificado como sendo K_i o sentido de P . O comportamento de dor de Job é um critério para determinar a existência da experiência de dor de Job, mas o termo «dor» não tem como sendo o seu sentido o comportamento de dor de Job. Segundo Wittgenstein é necessariamente verdadeiro que o comportamento de Job é evidência da sua experiência, mas evidentemente isto não é o mesmo que dizer que o termo «dor» significa o choro de Job. Pelo contrário, como vimos durante a exposição da doutrina da eliminabilidade da expressão de autoconhecimento, a expressão verbal substitui o choro, não é a sua descrição.

Foi já proposta a ideia de que, na filosofia da consciência, um critério não tem que ser necessariamente ele próprio observável. Assim propõe-se que um critério para determinar se eu compreendo a regra de Modus Ponens consiste no facto de se estar em condições de determinar se eu domino a regra. Mas dominar uma regra é possuir uma faculdade a qual ela própria não pode ser observada mas apenas inferida. O erro aqui consiste em tentar determinar o critério como sendo a faculdade. Como já foi dito, um processo interior não é um critério; ao contrário é para se estabelecer a sua existência que se tem que introduzir critérios. Assim, se a existência da faculdade se infere do seu exercício, é o seu exercício que tem que ser considerado como critério e portanto é observável.

Igualmente proeminente é o uso da relação de critério entre a fórmula K_i e uma proposição P durante a demonstração do argumento contra a existência da linguagem L_p . Como vimos, Wittgenstein argumenta que é impossível proporcionar uma definição ostensiva privada de uma sensação S , porque uma definição $D_t(S)$, i.e., a definição da sensação S no momento t , para ser adequada tem que satisfazer a condição $D_t(S)$ ser equivalente a, para qualquer k , $D_{t \cup k}(S)$. Mas é evidentemente que se S é um acontecimento privado e incommunicável, não existe um critério K_i que garantira a associação de « S » com a sensação original. A memória do sujeito não pode funcionar como sendo K_i uma vez que cada recordação necessita por sua vez de um critério para a legitimar como correcta e cairíamos assim num *regressus ad infinitum*. Por outro lado, quando o que há a recordar é um objecto privado, não há possibilidade de distinguir uma memória exacta de uma memória ilusória. Logo, a definição privada ostensiva $D_t(S)$ não pode ser efectuada.

Assim parece ser possível deduzir do argumento contra a existência de L_p que um conceito C é vazio se não existe um critério K_i tal que (K_i, C) . K_i tem que ser concebido pois como evidência não indutiva para a legítima aplicação de C . No Capítulo 3 de *Espontaneidade da Razão* discutiremos alguns dos aspectos do uso desta ideia em Lógica.

Wittgenstein não pretende afirmar que um conceito C que tem um critério K_i é sempre empregue com base em K_i . Pelo contrário, os conceitos de dor, de imagem mental, por exemplo, têm critérios que no entanto não são empregues por mim quando falo acerca de minhas dores e das imagens mentais que tenho neste momento na minha consciência. Uma observação, feita na terceira pessoa do indicativo, difere de uma asserção feita na primeira pessoa do indicativo acerca de uma impressão sensorial, pelo facto de aquela poder ser feita sem o recurso a qualquer inferência; difere ainda da asserção feita na primeira pessoa por só poder ser legitimada pelo recuso à evidência e , assim, a um dado K_i . É claro que K_i pode ser ele próprio uma impressão sensorial caso em que o sujeito não o pode reconhecer através de, digamos, K_j . No entanto os conceitos usados para falar acerca de K_i , K_j , etc., têm que ser eles próprios legitimados por novos critérios.

Parece ser seguro representar a posição de Wittgenstein como sendo a de que, supondo assegurada a igualdade entre critério e evidência conclusiva, uma modificação num critério K_i para um conceito C implica uma modificação no sentido de C . Um processo para determinar se K_i é um critério para uma proposição P consiste em ver se é possível descobrir empiricamente que K_i não é evidência a favor de P . Assim não é possível separar K_i de C sem alterar o sentido de C . Ao contrário, como já foi dito, uma vez que K_i não é o único critério para C , tem-se que é concebível juntar a K_i critérios K_j , K_k , ..., etc., para C sem alterar o sentido de C .

M. S. Lourenço
Universidade de Lisboa

Book reviews

Ezbozo de la Filosofía de Kripke, by Manuel Pérez Otero. Barcelona: Montesinos, 2006, 276 pp.

Manuel Pérez Otero's *Ezbozo de la Filosofía de Kripke* is a brief monograph on Saul Kripke's *Naming and Necessity* (Cambridge (1980): Harvard University Press), the book that was based on a transcript of the famous three lectures that Kripke gave at Princeton University in 1970. *Naming and Necessity* is one of the most discussed and influential books of 20th century analytic philosophy. Here and in a few other articles published between 1970 and 1980 (among which 'A Puzzle about Belief,' in *Meaning and Use*, ed. by A. Margalit, Reidel: Dordrecht and Boston) Kripke established the agenda for the subsequent investigations in philosophy of language, metaphysics and epistemology.

Two main paradigms characterize the debates in philosophy of language during the 20th century. The first is what Kripke calls the Frege-Russell theory of proper names that says that a proper name means the same as a certain definite description. The second was initiated by Ruth Barcan Marcus ('Modalities and Intentional Languages,' *Synthese*, 13, 1961, pp. 303–22) who termed proper names as mere 'tags' and it was later promoted by the contributions of Kripke, Keith Donnellan and Hilary Putnam at the beginning of the '70s and by many others since then. Their proposals differ in many aspects but some philosophers have been tempted to assimilate them into one and the same theory. For instance Quentin Smith ('Marcus, Kripke, and the Origin of The New Theory of Reference,' *Synthese*, Vol. 104, No. 2, August 1995, pp. 179–89) identified Marcus's view of proper names with the theory of reference given in *Naming and Necessity*. However, the idea of proper names as mere tags without descriptive content that is so typical of Millianism is not Kripke's idea of names as rigid designators, i.e. names that designate the same object with respect to all relevant worlds, for 'rigidified' definite descriptions are rigid designators but still have a descriptive content.

Among the many merits of Pérez Otero's monograph is the fact that he carries out an impressive work of systematic clarification of the fundamental aspects of *Naming and Necessity* by carefully avoiding typical,

mistaken generalizations and yet not requiring any previous specialist knowledge by the reader. For these reasons, I think that his monograph can be indicated as an appropriate guide for any graduate course.

The book is structured in three parts. The first part (Ch. 1–2) defines the philosophical context in which *Naming and Necessity* was written and the fundamental traits of the Frege-Russell theory of language that Kripke criticizes. The second part (Ch. 3) presents Kripke's objections to this paradigm and to its complementary metaphysical and epistemic theses. Furthermore the chapter presents Kripke's positive theses about necessity and its relations to epistemic concepts, the notion of possible worlds, his endorsement of Aristotelian essentialism, the relation between a name and its denotation, the nature of particular objects and of natural kinds. Three main chapters (Ch. 4–6), which are intended as three appendixes, constitute the last part. They are dedicated to Kripke's arguments supporting mind-body dualism (Ch. 4), his further criticisms to the Frege-Russell theory of belief reports (Ch.5) and some general conclusions (Ch.6).

Here I will concentrate on some fundamental matters that Pérez Otero treats in a particularly illuminating and original way. I start with his acknowledgement of John Locke's general empiricism as the proper source of many theses of the Frege-Russell paradigm.

Some of the upholders of this paradigm — among which Bertrand Russell, Rudolf Carnap and W.O. Quine — explicitly recognized the influence of David Hume's *Enquiry Concerning Human Understanding* (1748) on their theories. But in the first chapter of his monograph Pérez Otero individuates in John Locke's *An Essay on Human Understanding* (1690) the real theoretical framework of such views (for a similar interpretation see: Manuel García-Carpintero, *Las palabras, las ideas y las cosas*, Barcelona (1996): Ariel).

Locke contends that *intuitively* we would say that the meaning of natural kind terms as 'gold,' 'water' or 'tiger' consists in a semantic relation to something extra-linguistic, a substance (gold or water) or an animal species (tiger). The *relata* of such semantic relations have observable external properties (for instance water is a liquid that has no colour, no taste or odour, it flows in rivers and seas, etc.), but they also have some other properties that are not directly observable (for instance, nowadays we know that water is H₂O). Properties of the last kind constitute what Locke calls the *real essence* of an object, i.e. an internal structure epistemically inaccessible for us that causally determines the external properties that we can perceive. These

external properties constitute what Locke calls the *nominal essence* of a natural kind, even though the term is ambiguous and sometimes Locke uses it to define the ideas produced in us by their perception. However, the nominal essence cannot determine the real essence of the correspondent natural kind because different people perceive different nominal essences and nominal essences are defined on the basis of their effects on people's minds. Hence, Locke maintains an anti-essentialist view with respect to natural kinds.

Pérez Otero individuates three other main theses (p. 35, p. 178–9) that together with Locke's anti-essentialism constitute the philosophical background of the Frege-Russell paradigm criticized by Kripke:

- (i) *Knowledge entails Certainty*: we can only know that of which we have certainty.
- (ii) *Principle of Transparency of Representational Content*: we know exactly the content of our representations, we know what we think just as we know what the words of our language mean.
- (iii) *Semantic Internalism (or Individualism)*: the meanings of the words of our language consist of subjective mental entities that can vary from individual to individual.

Thesis (i) has been rejected by contemporary epistemology. Thesis (ii) is implicit in the Fregean conception of sense and in Russell's principle of acquaintance. However, one cannot apply an analogous principle at the level of denotation. In fact it is possible that we do not know that 'Aristotle' and 'the Stagirite teacher of Alexander the Great' share their referent. Hence, it seems that meaning should be at the level of sense rather than at the level of reference. And thesis (iii) was still endorsed by Russell, but it was already criticized by Frege's notion of sense as intersubjective and publicly available to the speakers of each linguistic community.

In Ch. 3 Pérez Otero shows how Kripke criticizes these views. First, let us consider Kripke's rejection of the Frege-Russell anti-essentialism that was originally proposed by Locke and later endorsed by Russell, Quine, Wittgenstein and Carnap among others. Kripke argues against them by appealing to our intuitive answers to certain mental experiments about modal properties. As Pérez Otero notices, this methodology is perfectly analogous to the scientific methodology. Scientific experiments try to build an ideal simplified situation; men-

tal experiments (through imagination) create a simplified situation in which we test our intuitions. In philosophy, intuitions have the same role that perception has in scientific experiments.

In the Frege-Russell tradition all necessity is linguistic or representational; it is a sub-product of the use of linguistic representational systems. If modality does not depend on the world, if objects do not have modal properties in themselves, then modality is reducible to the epistemic-representational domain and the truth-value of modal statements depends on language itself or on the internal states of our concepts. Kripke relates the epistemic-representational theory of modal discourse to descriptivism coherently to Quine's own distinction between *de re* and *de dicto* modality ('Reference and Modality,' in *From a Logical Point of View*, New York (1961): Harper & Row). But if following Quine we assume that necessity is predicated primarily of sentences (namely, if we take *de dicto* necessity to be fundamental) it is impossible to derive a concept of necessity predicated of the things themselves (namely *de re* necessity) independently of how we represent them.

Metaphysical necessity instead is predicated primarily of the relation between properties and particulars. From here, you can derive an application to representations. Essentialism agrees with our pre-theoretical conception of the world. So, it is the anti-essentialist theorist who has to prove that essentialism is wrong. The expressions 'essential properties' and 'essentialism' express concepts and theses that pertain to common sense. An essence is what characterizes univocally an entity and can be thought as the collection of all the essential properties of such an entity that together distinguish it from others. For instance, water is H₂O, and being H₂O is its essence because no other entity different from water is H₂O. It is more controversial whether or not particular objects have an essence, namely a set of properties that together distinguish (or identify) an object from any other. For instance, Socrates is essentially human, but *being human* is a property shared by all human beings. This means that the fact that an object has essential properties is not sufficient to individuate its essence.

An important objection to essentialist philosophers is that they cannot offer a general and coherent conception of essential properties. But Kripke and David Wiggins (*Sameness and Substance*, Oxford (1980): Blackwell) respond to this objection. The answer is in the distinction between the modal profile of rigid designators and definite descriptions. Saying that 9 is an odd number (a necessary truth) is

neither synonymous nor equivalent to saying that *the number of planets is an odd number* (a contingent truth).

The second line of criticism to the Frege-Russell paradigm rejects thesis (ii). Kripke endorses the kind of semantic externalism that Putnam and Donnellan in those same years were independently promoting. This is the thesis that the meaning of some signs of our language (for instance proper names) depends on external entities that do not reside in the minds of each individual speaker. For Kripke (and for Donnellan and Kaplan) this thesis coincides with the causal theory of reference that explains the connection between a proper name and its referent in terms of causal relations. A name can be introduced into the public language by fixing its referent either by ostension (through a direct perceptual relation) or by using a definite description (a postulation) as 'Let us use 'Julius' for whoever invented the zip.' Once introduced, the name can be transmitted from one speaker to another through chains of communication. As is well known Gareth Evans ('The Causal Theory of Names,' in *The Philosophy of Language*, ed. A. P. Martinich, (1985): Oxford University Press) criticized the causal theory because it does not produce sufficient conditions to the determination of the referential relation between a proper name and its referent. It is therefore a modest theory. Furthermore it does not allow establishing a link between the name and its referent without using the notion of reference, and hence it does not offer a reductionist analysis of the concept of reference. However, Pérez Otero (p. 176) notices that Kripke does not aspire to a different proposal and points out two main reasons for the definitional non-eliminability of the concept of reference. The first is that the theory indicates only necessary and not sufficient conditions for the correct application of the schema 'z is the referent of the proper name 'X'.' The second is that you cannot have a reductionist definition of 'z is the referent of the proper name 'X',' because the causal theory invokes the concept of reference both in the mechanism of introduction and in the mechanism of transmission of the name.

However, Pérez Otero (see section 6.2) also notices that there are a few important distinctions concerning Kripke's attitudes towards the principle of transparency. It seems that in 'A Puzzle about Belief' Kripke rejects it completely by sustaining that the meaning of a rigid designator is exhausted by its referent (here he favours a strict Millian position). But in *Naming and Necessity* things are not yet so radical. In fact, here the notion of meaning that he endorses is richer and in-

cludes both the referent of a rigid designator and its associated descriptive content. Hence, one might think that Kripke does not completely reject (ii), because (ii) might be assumed at the level of descriptive content and rejected at the level of denotation.

But Pérez Otero also notices that in *Naming and Necessity* Kripke puts forward another thesis that can be distinguished from his semantic externalism, what Pérez Otero calls *linguistic* or *semantic particularism* (see sections 1.1, 3.4, 6.2). This thesis consists in recognizing that rigid designators (among which names) have a singularizing function that is external to the mind of each speaker, that integrates the communicative chains and that works also when these terms appear in modal statements. These singular terms, irreducible to the predicative function of predicates and descriptions, are the basic building blocks of our language. One can fairly say that in *Naming and Necessity* Kripke substitutes the individualistic thesis of the principle of transparency into a social account of the meaning of our words. The relations with other subjects of the linguistic community are fundamental to the determination of linguistic meaning. Language, coherently with what the later Wittgenstein contends in *Philosophical Investigations*, is social.

I concentrated on what I think to be the most original theses of Pérez Otero's analysis of *Naming and Necessity*, but many more details on Kripke's arguments are to be found in the book. I intentionally focused on Pérez Otero's characterization of the Frege-Russell theoretical framework and on its origins in Locke's empiricism to point out the merits of his monograph as a rich, complex and yet coherent analysis of a difficult text whose many theses have rarely received such a systematic work of interpretation and clarification.

Fiora Salis

LOGOS — Logic, Language and Cognition Research Group
Departament de Lògica, Història i Filosofia de la Ciència
Universitat de Barcelona
fiora.salis@gmail.com

The Nature of Normativity, by Ralph Wedgwood. Oxford: Oxford University Press, 2007, 296 pp.

Ralph Wedgwood's *The Nature of Normativity* is a comprehensive book that aims simultaneously at providing a meta-ethical account of normative properties and statements and a discussion of the relevance of such a theory to issues pertaining to epistemology, philosophy of mind, metaphysics, and theory of rationality. His project is clearly stated from the beginning of the book. The view offered can be regarded as a variety of meta-ethical realism, which he claims to be compatible with a non-reductive understanding of normative properties and able to explain Normative Judgement Internalism (NJI) — the claim that there is an essential internal connection between normative judgement and practical reasoning or motivation. To do so he tries to offer a semantic account of normative concepts based upon conceptual role semantics — which he opposes both to causal accounts and to conceptual analysis accounts — and the underlying idea that the intentional is normative.

Before defending his own view, Wedgwood sets up a clear map of the views in dispute paying special attention to those aspects of his own project that cannot be successfully explained within those rival accounts. Hence the discussion he establishes with rival views always have a foot on the points he takes as most significant for his own project. Thus, for example, Wedgwood assumes NJI is true and focuses upon the impossibility of explaining this claim within Expressivism, Causal Theories and Conceptual Analysis accounts. For him, the essential connection between normative judgments and practical deliberation can be better explained if we take the former as expressing cognitive states — i.e. beliefs. He needs to argue, against those who defend a Humean account of motivation, that mere beliefs are sufficient to make NJI true. To do so, Wedgwood introduces what might be taken as the core of his proposal: an analysis of the concept 'ought' in terms of conceptual role analysis.

According to this approach, the nature of a concept is the essential conceptual role it plays in reasoning. For example the conceptual role played by the logical operator 'or' is given by the basic rules of rationality that governs its use, i.e., that make 'if p, then p or q ' valid. Thus, someone is said to possess a concept if she manifests a disposition to follow the basic rules of rationality for using it. Specifying

those basic rules will give us an explanation of what it is to possess a particular concept and the semantic value of that concept.

In the case of normative concepts, such as 'ought,' that role consists, according to Wedgwood, in a certain regulative role the concept plays in *practical* reasoning. So, it is part of the basic rational rules of the use of the concept that there is a relationship between using the concept and some facts about motivation.

To establish this, Wedgwood cannot simply take for granted that the concept 'ought' plays this essential conceptual role in practical reasoning, for it can be disputed whether a concept might have this sort of role at all. Wedgwood needs then to articulate that particular conceptual role in such a way that it follows from the basic rules that govern its rational use that it possesses this connection to motivation or practical reasoning. As he explicitly puts it, 'Acceptance of the first-person proposition 'O <me, t > (p)' — where ' t ' refers to some time in the present or near future — commits one to making p part of one's ideal plan about what to do at t ' (p. 97). According to this characterization, the rational use of the concept 'ought' is based upon my disposition to follow the following rule: my recognition that I ought to ϕ commits me to adopt as part of my ideal plan any p that is implied by ϕ . Not following this rule will amount to not using the concept 'ought' in a rational way. One can ask at this point in virtue of what can Wedgwood establish that this is the rule that governs the rational use of the concept. Why will I be irrational if I use the concept 'ought' in a way that does not fit with this rule? To some extent the simplest answer is that there is nothing more basic one can appeal to in order to justify the rule. According to Wedgwood, this rule is basic and users do not acquire it in any sort of inferential or justificatory way. Put crudely, rationally using the concept amounts to having the disposition to use it according to this rule.

At some point, Wedgwood notices that accomplishing these rules is not simply a matter of conforming our practice to them. Thus, following his analogy with belief states, there is a point in conforming to the standards of justification and rationality of beliefs. Hence, the point of achieving correct and rational beliefs is that we 'get things right' regarding facts about the world. Similarly, Wedgwood holds, there is a point in adopting correct and rational plans: to get things right in one's plans and intentions and hence 'have a set of intentions that one will actually execute in such a way that as a result one will act in a manner that is genuinely choiceworthy.' (p. 101) Although Wedgwood's

analogy between the correctness of a plan and a belief might seem persuasive, there is a certain ambiguity in the way he characterizes the conditions under which a plan is correct. On the one hand, he seems to claim that the correctness of an ideal plan depends upon external facts about the agent (Wedgwood's, pp. 101–102) but, on the other, it seems that a plan is correct if one would act in a genuine choiceworthy way, which, in turn, implies that one will perform p . It is in virtue of this second characterization that Wedgwood can dispel a possible objection to his characterization of the essential internal connection between ought and practical rationality. According to this objection, it is possible to find a case where one ought to plan on an action but one does not have to take that course of action. The example is of an eccentric millionaire who is willing to give you £1 million if you plan to drink a toxin but who is quite indifferent to the fact that you actually drink it. Wedgwood dispels the objection by claiming this is not a case where the agent forms a *correct* plan in the relevant sense, because 'for a plan to be correct, it would have to be true that if you act in a genuinely choiceworthy way at t , then you will drink the toxin at t .' (Wedgwood, p. 104) I think one can be quite unconvinced by this answer. Why is the agent's decision to plan on drinking the toxin, but not actually carrying on the plan, not a choiceworthy action? His answer seems to simply restate the conditions that the example challenges, but this does not seem to be a good answer. It seems that it would be correct to plan on doing p if p is part of a correct plan, but then a correct plan cannot be simply characterized in terms of that plan that will commit me to perform p .

The notion of a correct plan is taken under more consideration in Part II of the book, devoted to the metaphysics of normative facts. In chapter 7, Wedgwood tries to clarify what it is for a plan to be correct. To do so, he will appeal to the general idea underlying his project that the intentional is normative. Thus, normative judgments possess propositional content and they are winning judgements iff the proposition embedded is true. Mental states have, therefore, both standards of justification and rationality; besides there is a purpose or goal in conforming to these standards. Wedgwood keeps using his analogy between belief and planning. To that extent, he draws a similar story for what counts for a belief to be correct and rational and what counts for a plan to be correct and rational. Since the final purpose of belief is to get things right about the world, a belief would be correct if it is true and rational if, 'in relation to a given body of

information I, that body of information I makes it *highly likely* that my belief in question is correct' (p. 156). Similarly, since the final goal or purpose of a plan is to make correct choices, a plan will be correct if 'one will realize those choices and act in a way that is genuinely choice worthy' and it is rational 'just in case given I, that choice maximizes one's rational expectation of coming as close as possible to choosing correctly.' (p. 162). Here, as well, the notion of a plan's being correct seems to be defined in terms that leave some room for ambiguity about what counts as a correct plan. Acting in a way that is genuinely choice worthy does not seem to be a property of the content of my plan but of the action that would allegedly follow if I adopt *p* as my plan. It might be true that the final goal or purpose of practical deliberation is to get things right about our plans or choices, but it seems we still need some more clarification of what counts as a correct plan if Wedgwood's thesis is to have the convincing character of the corresponding thesis for the case of belief.

In the rest chapters of part II, Wedgwood discusses some of the implications of the claim that the intentional is normative for some metaphysical debates. In particular, he tries to show how this claim implies that normative properties cannot be reduced to non-normative ones, although he thinks this is still compatible with a form of naturalism.

In part III of the book, Wedgwood addresses some epistemological aspects of his view. Since he has committed himself with metaphysical realism about normative properties, he will endorse a form of cognitivism. Once more, he introduces an analogy with the acquisition of perceptual beliefs in order to delineate his own epistemology of normative beliefs. He claims our knowledge of cognitive statements requires us to acknowledge that there are 'primitively rational ways of forming beliefs' (p. 230) which we take as basic sources to acquire those beliefs and that cannot be justified in any further way. Allegedly, this notion is to play some role in our acquisition of normative beliefs as well. However, there seems to be a problem with the way in which Wedgwood establishes this notion. The problem can be easily shown in the case of belief, but I take it can also be of relevance in the case of normative beliefs. For example, in the case of perceptual beliefs, Wedgwood claims one can rely on perceptual experience as one of the basic ways of forming beliefs because 'There *is* an essential connection between our sensory experiences and the truth' (p. 232) However, this can only be stated in the absence of the hypothesis of

the evil demon. In that case, there would not be an essential connection between our sensory experiences and the truth. Therefore, if anything like a 'primitively rational way of forming beliefs' is at stake in acquiring normative beliefs, a similar worry might arise regarding whatever that way is thought to be. In particular, since Wedgwood claims that way is a form of normative intuition, exercised and manifested in our dispositions to respond to normative facts in the corresponding ways, we may have worries about the essential connection between our intuitions and truth.

Finally, Wedgwood tries to defend the claim that normative knowledge is *a priori* and he further discusses how his account is able to accommodate genuine normative disputes.

Certainly, any reader interested in meta-ethics will find in this book a very rich and comprehensive attempt to delineate the nature of normative statements. Although there Wedgwood offers very appealing claims, there seem to be a few explanatory gaps within some of the points discussed that may leave the reader uncertain regarding crucial aspects of his view.

María José Alcaraz León

Dept. de Filosofía, Facultad de Filosofía, Universidad de Murcia
Edificio Luis Vives. Campus de Espinardo, 30100 Murcia, Spain
mariajo@um.es

When Truth Gives Out, by Mark Richard. Oxford: Oxford University Press, 2008, 184 pp.

Richard's *When Truth Gives Out*, written in an engaging and accessible style, develops around the idea that the notion of truth, contrary to a lot of received wisdom from philosophy of language and logic, is not — or at least, not always — the right concept to employ in analyzing belief, assertion, or their evaluation. The book is organized in five chapters and two appendices, all of which could work equally well as independent essays. In particular, Chapter IV, *What's the Matter with Relativism?*, largely overlaps with his well-known paper 'Contextualism and Relativism,' *Philosophical Studies* 119, 2004, 215–42. The choice of compiling those pieces into a monograph, rather than a mere collection of essays, is motivated by the fact that each chapter addresses, sometimes in different ways and from different angles, the

question of whether the notions of truth and falsity are the core notions in the analysis of a range of phenomena regarding thought, discourse, inference and disagreement. The take-home message of the book is that if one is prepared to give up the centrality of the notion of truth — whether one trades it for the notion of *relative* truth or, more radically, decides that even relative truth simply is not the right dimension to understand belief and assertion — then one can start perceiving promising solutions to a number of long-standing puzzles from philosophy of language, logic and epistemology.

Chapter I, *Epithets and Attitudes*, provides a fairly engaging introduction to the matter, with its topic of derogatory terms and, more specifically, racial slurs. Consider the somewhat *démodé* derogatory term for the French ‘Frog,’ and suppose that I point to someone and say ‘He is a Frog.’ What did I say? Did I say something true, or false, or neither? Which thought did I express (if any)? Chapter I deals with this sort of questions, aiming at the conclusion that truth-conditional approaches — i.e., approaches that take the notion of truth to be the central notion in the analysis of meaning — run into trouble when it comes to accounting for the meaning of slurs, and of derogatory terms more generally. Although no one holds that ‘Frog’ is simply *synonymous* with ‘French,’ a minimal departure from such a simple view would be to say that ‘He is a Frog’ has the same truth conditions as ‘He is French,’ but that the former, unlike the latter, also *presupposes* a derogatory attitude towards the French. Similar in spirit, albeit different in detail, is the proposal that while ‘Frog’ and ‘French’ are truth-conditionally equivalent — both predicates apply to the French, and to the French only — a speaker who uses ‘Frog’ rather than ‘French’ *conventionally implicates* that the French are worthy of contempt (or something along those lines). Presuppositional accounts, of the former sort, are discarded on the charge of ‘misdiagnos[ing] ‘how slurs work’ (p. 21): someone who is using ‘Frog’ is acting in a contemptuous and hostile way towards the French, rather than merely presupposing such a contemptuous attitude and ‘inserting [it] into a conversation’s record’ (p. 22), which is what Richard takes such presuppositional accounts to amount to. While this argument against presuppositional accounts may appear somewhat elusive and thin, Richard’s dissatisfaction with the ‘pragmatic’ accounts, of the latter sort, which locate the derogatory character of slurs at the level of conventional implicature, rather than at the level of truth-conditional or ‘semantic’ content, is more substantial. Richard’s worry is that

such accounts (and, *ipso facto*, presuppositional accounts, too) fail to look beyond the language and its use. But slurs are not merely a linguistic phenomenon: they reach beyond language into the thought itself. A person using a racial slur is not merely implicating something negative: the thoughts of this person are also infested. The way in which slurs infiltrate the thought is, in Richard's view, a core problem that will elude any account whose analysis remains at the level of language. Richard's own account is thus characterized by the idea that the way in which the thinker *represents* the object that he or she is thinking about is constitutive of the thought itself. My thought that a certain person is a Frog would thus be a thought with a built-in contemptuous attitude towards the French, a feature that distinguishes it from the thought that the person at stake is (merely) French.

Chapter I offers a novel, albeit somewhat underdeveloped approach to slurs and derogatory terms, and, more importantly for the book's main focus, puts forward the idea that the analysis of thoughts expressed or entertained by a slurring agent overrides the dimensions of truth and falsity. The argument may be reconstructed as follows. An agent's contemptuous attitude towards the French is part of her thought that *X* is a Frog; to think such a thought is to think badly of *X* (and of the French more generally). Even if *X* is French, we cannot accept, or adhere or ascribe to, that agent's thought, because that would imply that we endorse her contemptuous attitude. On the other hand, we cannot unsubscribe from her thought either on the grounds on its being *false*, because that would commit us to endorsing that very same contemptuous attitude: to ascribe *falsity* to that agent's thought is to ascribe *truth* to the thought that *X is not* a Frog; and to think that someone is not a Frog is, again, to think badly of the French. *Ergo*, the thought of a slurring agent can neither be true nor false. Truth gives out.

In Chapters II and III, Richard turns to a different set of problems. He considers, on the one hand, the phenomenon of *vagueness*, and on the other, the *liar paradox*. What is the connection between (complex) expressions such as the liar sentence, namely 'This very sentence is not true' and vague expressions such as 'bald,' 'tall' or 'many'? Why should the two issues be grouped together and treated along the same lines? Although Richard does not try to articulate the connection explicitly, the likely motivation is that in both cases, a sentence that (arguably) cannot be true still seems to 'get things right.' The liar sentence denies its own truth and thus provably

cannot true; yet precisely in denying its own truth, there is something that it gets right. Similarly, a *prima facie* contradictory sentence such as 'Joe is not bald and it is not the case that Joe is not bald,' in a situation in which Joe is a borderline case of baldness, also seems to get things right, yet one would not necessarily want to take it to be saying something *true*. So the puzzle is this: how can a person say something and be right without *ipso facto* asserting something true? Richard's answer to this question is that, beside assertion, which he thinks should be analyzed as commitment to *truth*, there are many other sorts of commitments. Elaborating on such other sorts of commitments, he puts forward a proposal whose central idea is that we can assess a commitment as *appropriate*, thereby accounting for the intuition that the speaker can, in some sense, be *right*, without having to invoke the notion of truth.

A long standing objection to views that, like Richard's, hold that not all speech acts are to be evaluated in terms of truth and falsity, is that they are unable to account for logical inference. Consider the inference 'Joe is not bald. Therefore, not everyone is bald.' Assume that 'bald' is a vague predicate, and let Joe be borderline bald. If, as Richard holds, the premise in the inference cannot be the assertion of the negation of Joe's being bald, but is, rather, a *sui generis* denial, how could he possibly account for the fact that the inference at stake is a *logically valid* inference? Does not validity have to be cashed out in terms of truth and falsity? Richard suggests that it does not, and offers an account on which sentences that serve as components (premises and conclusions) of a logically valid argument need not be 'vehicles of assertions,' but can be vehicles of various other sorts of commitments, of which *assertoric* commitments are only a subset. The trick is to analyze particles such as 'not' or 'if...then' as devices for compounding commitments *in general*. 'If Joe is not bald, then not everyone is bald' can, then, be viewed as not only a well-formed clause, but also as a logically valid one, even if neither the antecedent nor the consequent are truth-evaluable. Evaluation of truth, in Richard's account, is just a special case of a more general and more basic type of evaluation, namely evaluation of *appropriateness*. Even when *truth* gives out, as in the case of vague expressions, liar sentences, and, more widely, all kinds of normative and evaluative discourse, there is still a notion of validity to hold on to, based on the idea that non-assertoric commitments can still be evaluated as appropriate or inappropriate.

The issue of vagueness is taken up again in Chapter IV, although from a different angle. Here, Richard is concerned with the question of how there can be disagreement over the application of a vague predicate, yet disagreement in which neither party is, properly speaking, wrong. Richard's working example is Didi and Naomi's disagreement over the issue of whether Mary, who won a million-dollar lottery, is rich. Didi holds that she is rich, while Naomi, for whom a million dollars is not that much, holds that she is not, yet there seems to be no objective way of adjudicating between Didi and Naomi. Richard is aware that gradable adjectives are sensitive to contextual variations: if Didi says that Mary is rich in a conversation about life conditions in derelict suburbs of third world countries, while Naomi says that she is not rich in a conversation about Bill Gates and Donald Trump, then it becomes easy to account for the idea that both Didi and Naomi may be right. The gradable adjectives' dependence on the so-called *comparison classes* is no news, and can be assimilated, in a more or less straightforward way, to familiar forms of context-dependence, like indexicality. Thus e.g. if I say 'It's raining' in Paris, and you say 'It's not raining' in New York, we can obviously both speak truth. But Richard insists that the phenomenon that he is interested in is not dependence on a comparison class, and cannot be assimilated to any form of indexicality. For, he argues, when the comparison classes are different, we lose the intuition of disagreement, to the same extent that we have no such intuition if I say that it's raining and you say that it is not, but I am in Paris and you are in New York. Indeed, Naomi might well accept that Mary is rich relative to the population of suburbs in third world countries, and that, at the same time, she is not rich compared to Trump, Gates, etc. Richard's scenario is, by assumption, one in which 'the salient comparison class is the same in both cases (it is, say, New Yorkers)' (p. 93). The reason why we can fix the comparison class and still be unable to adjudicate between Didi and Naomi is that each places the threshold of what it takes to be rich at a different level of the richness scale. Naomi's threshold is below one million dollar, while Didi's is further up.

Now, gradable adjectives' dependence on a threshold point *in addition to* a comparison class (or a scale) is not really news either: the idea can be traced at least back to Hans Kamp, 'Two Theories of Adjectives,' in *Formal Semantics of Natural Language*, ed. by Ed Keenan, Cambridge University Press, Cambridge, 1975, 123–55. But, wheth-

er or not Richard, noticeably parsimonious in his references to the vast literature on vagueness and gradable adjectives, is aware that the importance of thresholds has been previously acknowledged, one should grant him the insight that thresholds and comparison classes, *qua* parameters on which truth value depends, are not on a par. While Naomi may consistently say ‘Mary is rich for the population of derelict suburbs of third world countries, but is not rich for the millionaires,’ it would be odd for her to say ‘Mary is rich for a New Yorker relative to a threshold set below one million, but is not rich for a New Yorker relative to a threshold set further up.’ To put the same point differently, one can be tall for an ordinary person and not be tall for a basketball player, but one either is or is not tall for a basketball player. Whether one is tall for a basketball player requires, of course, fixing a cut-off point within the tallness scale (as restricted to the class of basketball players). But once the comparison class is fixed, one cannot go on, it seems, playing around with thresholds (or, at least, not with the same flexibility with which one can play around with comparison classes). So, summing up, when *F* is an adjective such as ‘tall’ or ‘rich,’ in order to interpret the claim that *F* applies to some object *X*, one must first determine the relevant comparison class. Parties who disagree whether *X* is *F*, while relying on different comparison classes, are only engaged in a spurious disagreement. This does not seem to hold for threshold-dependence. When Naomi and Didi disagree whether Mary is rich for a New Yorker, the interpretation is not at issue. Each understands what the other is claiming, there is no misunderstanding of any sort, and yet, we seem to lack objective grounds on which to adjudicate between them, to decide that the one rather than the other got the facts right. In a disagreement in which neither party is at fault, again, truth gives out.

But how can there be genuine, substantive disagreement that is not cashed out in terms of truth and falsity? Richard’s answer is twofold. First, Didi and Naomi’s disagreement is substantive because they share all the concepts involved in the issue over which they disagree, and, in particular, the concept denoted by the predicate ‘rich.’ To illustrate the point, suppose that Mary is a dentist, and Didi says ‘Mary is good’ while Naomi says ‘Mary is not good.’ If Didi means to be saying that Mary is a good dentist, while Naomi means to be saying that Mary is not a (morally) good person, then their uses of the predicate ‘good’ arguably denote different concepts, and their disagreement is merely verbal, rather than substantive. Secondly, it is not

necessary, according to Richard, that either of the disagreeing parties should get the facts right and latch onto truth. In the picture that he proposes, there is a certain concept in place, the concept of being rich, but the *extension* of the concept, i.e. the range of individuals to which the concept applies, is not determined by the concept itself (together with the facts). Rather, it is open-ended, and gets more and more determined as the concept starts getting entrenched into a linguistic community. Richard thus writes: 'Because 'rich' is subject to accommodation, speakers in different conversations (...) are able to impose different extensions (and thus intensions) on 'rich'; the result is that different speaker's uses of 'Mary is rich' can have different truth-values' (p. 100); 'What a concept or notion is a concept or notion *of* gets worked out over time via something like a process of cultural accommodation and negotiation' (p. 116). In other words, accommodation and negotiation are processes that make room for a kind of disagreement that falls in between, on the one hand, merely verbal disagreement and, on the other, merely factual disagreement, hence a disagreement that is arguably substantial, yet overrides the dimension of purely extensional notions such as truth and falsity.

Richard's proposal in Chapter IV is interesting and plausible, though not without gaps. One such gap concerns the methodology. Are there any criteria to decide whether a given disagreement is a verbal issue, rather than a case of accommodation and negotiation? And are there any criteria to decide which aspects of meaning are open to accommodation and negotiation and which are not? While Richard does not lay down any such criteria, he implicitly uses some, as when he writes: 'when variation in an expression's semantics [is] not open to contextual negotiation, that variation affect[s] what is said' (p. 102). But this is a very shaky criterion. According to Richard, when Naomi says 'Mary is tall' in a context in which the relevant comparison class is her basketball team, what she really says is that Mary is tall *for a player of her basketball team*. But this does not seem correct: even in such a context, what Naomi says is, rather, that Mary is tall, and that's all. Another gap is that the issue of what it takes for two people to share a concept is never really addressed. Thus if Naomi considers only millionaires to be rich, why should we say that her concept is the same as Didi's concept 'rich'? The boundaries between the cases in which we have distinct, albeit related concepts, and those in which there is a single concept, subject to accommodation and negotiation, are as vague as the concepts under considera-

tion. Consequently, whether a disagreement is verbal or not, and whether it is substantive or not, are issues that cannot, and should not, be settled in some objective, context-independent way.

The previous point is not really hostile to Richard's overarching view. Indeed, the proposal that he presents in Chapter V, *Matters of Taste*, goes in the same direction. He is still concerned here with disagreement, but turns his attention to disagreements over what is cool, who is hot or good-looking, what food is yummy, and so on. In Richard's view, 'judgments about what is cool and other matters of taste enjoy a *double relativity*' (p. 132, my italics). Consider our disagreement on the issue of whether some person, e.g. Ethan Hawke, is attractive. We can both give reasons and arguments that support our respective judgments. Our disagreement is substantive, yet there seems to be no objective grounds to say that the one rather than the other got the facts right. From my perspective, Hawke is attractive, from yours, he is not, and the reason why it makes sense to continue arguing about the issue is that our shared concept of attractiveness is subject to accommodation and negotiation; we have, again, the same sort of relativity as the one discussed in Chapter IV concerning disagreement over who is or is not rich. But Richard now points out that 'not only may the assignment of a truth-value to a claim be 'perspective-relative,' but whether a claim is truth-apt may also be 'perspective-relative'' (p. 126). The idea is that in some cases, our respective evaluations of Hawke as attractive are not so much judgments that he is or is not attractive, as judgments (or reports, if you prefer) that we find him attractive or do not. While the proposal is compelling, it is somewhat unclear why it would only arise with judgments on matters of taste. Why should it be implausible, for instance, to hold that Didi's evaluation of Mary as rich might be of the same kind as my evaluation of Hawke as attractive? In either case, the agent would be simply *asserting her own perspective*, to use a phrase from Lasersohn (Peter Lasersohn, 'Context Dependence, Disagreement, and Predicates of Personal Taste,' *Linguistics and Philosophy* 28, 2005, 643–86), rather than making a truth-apt judgment to the effect that Mary *is* rich or that Hawke *is* attractive.

A more serious problem for Richard's account of disagreement, as well on issues involving concepts such as 'rich' as on matters of taste, is that, as he has acknowledged himself, 'all the hard work is yet to come' (p. 123). For, Richard insists that when a concept is subject to accommodation and negotiation, the disagreement is substantive, and

it makes sense for the disagreeing parties to go on disagreeing, to challenge one another's evidence and arguments for the claims that they are putting forward, to try to prove the other party to be wrong. In other words, the disagreement should allow for a rational resolution; at the end of the day, the process of accommodation and negotiation should result in having one of the parties emerge as the one that somehow 'got it right.' But if truth has given out from the start, what else is there to ground the one party's perspective as somehow better, or superior, or 'more right,' than the other's? Richard, while fully aware of the issue, has preferred not to address it in this book.

When Truth Gives Out is written in an entertaining, almost colloquial tone, with most of the technicalities confined to appendices, and dry argumentation carefully avoided. However, this 'easy' style can easily prove to be treacherous. For, the thread of argument is often elusive, and the precise reconstruction of the argument may require considerable effort. What is more, the discussion is often sparkled with interesting and provocative ideas, but whose role within the general argument is not always clear. Thus, for example, in Chapter I, we find the idea that 'whether a thought is true or false can be as much a matter of the context in which it is asked *Is that true?* as it is of how things are independently of asking' (p. 15). The idea is deeply 'relativist': the truth of a thought is not fixed once and for all by what the world is like, but is relative to the context from which the thought is evaluated for truth; the same thought, evaluated as true in one context, may be (correctly) re-evaluated as false from another. However, the exact contribution of this relativist stance to the discussion of derogatory terms and racial epithets remains unclear. Throughout the book, one often gets the impression of being presented with a patchwork of interesting, thought-provoking ideas, rather than a single, unified and well-structured picture. But despite this somewhat impressionistic character, *When Truth Gives Out* is rhythmmed by a unifying motto, aptly captured by its witty title.

Isidora Stojanovic
Institut Jean Nicod, CNRS-ENS-EHESS
isidora.stojanovic@ens.fr