



PENSAMENTO E FALA^a

Donald Davidson

Tradução de Marcelo Fischborn

Mestrando pela Universidade Federal de Santa Maria (bolsista CAPES) – RS - Brasil
marcelo_fischborn@yahoo.com.br

Qual é a ligação entre pensamento e linguagem? A dependência do falar em relação ao pensar é evidente, pois falar é expressar pensamentos. Essa dependência manifesta-se de infindáveis outras maneiras. Alguém que profere a frase “A vela está apagada” como uma frase do português tem de ter a intenção de proferir palavras que são verdadeiras se, e somente se, uma vela indicada está apagada no momento do proferimento, e tem de acreditar que, produzindo os sons que produz, estará proferindo palavras que são verdadeiras apenas naquelas circunstâncias. Deter-se nessas intenções e crenças não é próprio do falante fluente. Mas, embora possa normalmente não chamar atenção, sua ausência seria suficiente para mostrar que ele não está falando português, e a ausência de quaisquer pensamentos análogos mostraria que não estava sequer falando.

A questão repousa no outro lado: pode haver pensamento sem fala? Uma reação inicial e natural é que pode haver. Existe a experiência familiar e incômoda de não se conseguir encontrar as palavras para expressar as ideias que se tem. Por vezes, pode-se julgar que o escritor de um editorial apresentou algo melhor do que nós próprios poderíamos fazê-lo. E há o cão de Norman Malcolm que, tendo perseguido um esquilo pelo bosque, late diante da árvore errada.¹ Não é difícil creditar ao cão a crença de que o esquilo está naquela árvore.

Uma intuição precisa, ainda que mais fraca, sugere o contrário. Pode-se perguntar se o falante que não consegue encontrar as palavras corretas tem uma ideia clara. Atribuir intenções e crenças a cães cheira a antropomorfismo. Um behaviorismo primitivo, aturdido pela privacidade de pensamentos não ditos, pode reconfortar-se com a concepção de que pensar é na verdade “falar para si mesmo” – discurso silencioso.

^a Publicado originalmente em *Mind and Language*, editado por Samuel Guttenplan (Oxford: Oxford University Press, 1975); reimpresso em Donald Davidson, *Inquiries into truth and interpretation* (Oxford: Oxford University Press, 1984). Nesta tradução foi usada a versão reimpressa em 1984. Agradeço a Rogério Passos Severo pela revisão técnica e a Marcia Cavell pela autorização desta tradução.

¹ N. Malcolm, “Thoughtless Brutes” (in: *Proceedings and Addresses of the American Philosophical Association*, 1972-3).

Sob a superfície dessas tendências opostas, operam correntes fortes, ainda que túrgidas, que podem ajudar a explicar por que os filósofos têm, no mais das vezes, preferido tomar uma posição sobre a questão a produzir um argumento. Qualquer que seja a razão, a questão da relação entre pensamento e fala parece raramente ter sido levantada por interesse nela mesma. A suposição usual é que um ou outro, fala ou pensamento, é comparativamente fácil de entender, e, portanto, o mais obscuro dos dois (qualquer que seja) pode ser esclarecido por análise ou explicação em termos do outro.

A suposição, penso eu, é falsa: nem linguagem, nem pensamento podem ser completamente explicados em termos do outro, e nenhum tem prioridade conceitual sobre o outro. Os dois estão certamente ligados, no sentido de que cada um exige o outro a fim de ser entendido; mas a ligação não é tão completa que qualquer um dos dois baste, mesmo que razoavelmente reforçado, para explicar o outro. Para sustentar essa afirmação, é necessário, sobretudo mostrar como o pensamento depende da fala, e essa é a tese que quero refinar para, na sequência, argumentar em seu favor.

Atribuímos um pensamento a uma criatura sempre que assertivamente empregamos uma frase positiva, cujo verbo principal é psicológico – em português, “acredita”, “sabe”, “espera”, “deseja”, “pensa”, “teme”, são exemplos – seguido de uma frase e precedido do nome ou descrição da criatura. (Um “que” pode, opcional ou necessariamente, seguir o verbo.) Algumas dessas frases atribuem estados, outras relatam eventos ou processos: “acredita”, “pensa” e “quer” relatam estados, enquanto que “veio a crer”, “esqueceu”, “concluiu”, “notou”, “está provando” relatam eventos ou processos. Frases que podem ser usadas para atribuir um pensamento exibem o que é frequentemente chamado, ou analisado como, intencionalidade semântica, que significa que a atribuição pode mudar de verdadeira para falsa, ou de falsa para verdadeira, por meio de substituições nas frases componentes que não alterariam o valor de verdade da frase isolada.

Não presumo que se uma criatura tem um pensamento, então podemos corretamente, com recursos do tipo recém-esboçado, atribuir esse pensamento a ela. Mas, ao menos, pensamentos assim atribuíveis constituem uma boa amostra da totalidade.

É duvidoso se os vários tipos de pensamentos podem ser reduzidos a um só, ou mesmo a alguns poucos: desejo, conhecimento, crença, temor, interesse, para nomear alguns casos importantes, são provavelmente independentes logicamente, à medida que nenhum pode ser definido usando os outros, ainda que com noções adicionais, tais como verdade e causa. Apesar disso, a crença é central para todos os tipos de pensamento. Se alguém está contente porque, ou nota que, ou lembra que, ou sabe que, o revólver está carregado, então precisa acreditar que o revólver está carregado. Mesmo para perguntar-se se o revólver está carregado, ou especular sobre a possibilidade do revólver estar carregado, exige-se a crença de que, por exemplo, um revólver é uma arma, que é um objeto físico mais ou menos duradouro, e assim por diante. Há boas razões para não insistir em qualquer lista particular de crenças que são necessárias se for para uma criatura perguntar-se se um revólver está carregado. No entanto, é necessário que haja intermináveis crenças enganchadas umas nas outras. O sistema dessas crenças identifica um pensamento situando-o em um espaço lógico e epistêmico.

Ter um pensamento exige que haja um pano de fundo de crenças, mas ter um pensamento particular não depende do estado de crença com respeito a esse mesmo pensamento. Se cogito de ir a um certo concerto, sei que terei algum grau de transtornos e gastos, e tenho crenças mais complexas sobre o prazer que experienciarei. Vou deliciar-me ouvindo, digamos, a *Grosse Fuge* de Beethoven, mas somente contanto que a execução atinja um padrão razoável e eu consiga permanecer atento. Tenho o pensamento de ir ao concerto, mas, até decidir que vou, não firmei a crença de que irei; até então meramente cogito o pensamento.

Podemos dizer, resumindo os dois últimos parágrafos, que um pensamento é definido por um sistema de crenças, mas é ele mesmo autônomo com respeito à crença.

Usualmente pensamos que ter uma linguagem consiste em grande medida em ser capaz de falar, mas, no que se segue, falar desempenhará apenas um papel indireto. O que é essencial ao meu argumento é a ideia de um intérprete, alguém que entende os proferimentos de outrem. As considerações a serem apresentadas implicam, penso eu, que um falante deve ser ele próprio intérprete de outros, mas não tentarei demonstrar que um intérprete tem de ser um falante, ainda que possa haver boas razões para sustentar isso. Talvez valha a pena assinalar que a noção de uma linguagem, ou de duas pessoas falando a mesma linguagem, parece não ser necessária aqui. Dois falantes podem interpretar cada um os proferimentos do outro sem que haja, em qualquer sentido ordinário, uma linguagem em comum. (Não quero negar que em outros contextos a noção de uma linguagem compartilhada pode ser muito importante.)

A tese principal deste artigo é que uma criatura não pode ter pensamentos a menos que seja intérprete da fala de outra. Essa tese não implica a possibilidade de redução, behaviorista ou outra, do pensamento à fala. Certamente a tese não imputa nenhuma prioridade à linguagem, epistemológica ou conceitual. A afirmação também fica aquém de afirmações similares, na medida em que permite que possa haver pensamentos para os quais o falante não consiga encontrar palavras, ou para os quais não haja palavras.

Alguém que é capaz de interpretar um proferimento da frase do português “O revólver está carregado” precisa ter muitas crenças, e essas têm de ser muito parecidas com aquelas que alguém precisa ter se cogita o pensamento de que o revólver está carregado. O intérprete precisa acreditar, podemos supor, que o revólver é uma arma, e que é um objeto físico mais ou menos duradouro. Provavelmente não há uma lista definida de coisas nas quais alguém que entende a frase “O revólver está carregado” precisa acreditar, mas é necessário que haja intermináveis crenças encadeadas umas às outras.

Um intérprete sabe as condições sob as quais proferimentos de frases são verdadeiros e, frequentemente, sabe que se certas frases são verdadeiras, outras tem de sê-lo. Por exemplo, um intérprete do português sabe que se “O revólver está carregado e a porta está chaveada” é verdadeira, então “A porta está chaveada” é verdadeira. As frases de uma linguagem têm um lugar no espaço lógico criado pelo padrão dessas relações. Obviamente o padrão de relações entre frases é muito parecido com o padrão de relações entre pensamentos. Esse fato encorajou a concepção segundo a qual é redundante tomar como básicos os dois padrões. Se pensamentos são primitivos, uma linguagem parece não servir a propósito algum que não expressar ou transmitir pensamentos; enquanto que se tomamos a fala

como primitiva, é tentador analisar pensamentos como disposições de fala: como diz Sellars, “[...] pensar no nível distintivamente humano [...] é essencialmente uma atividade verbal”.² Mas claramente o paralelo entre a estrutura de pensamentos e a estrutura de frases não fornece nenhum argumento para a primazia de qualquer um, e, sim, apenas uma presunção em favor da sua interdependência.

Estivemos falando livremente de pensamentos, crenças, significados e interpretações, ou melhor, usando livremente frases que contêm essas palavras. Mas certamente não é claro que entidades, ou tipo de entidades, têm de haver para que essas frases façam sistematicamente sentido. No entanto, aparentemente falar de pensamentos e dizeres de fato pertence a um modo familiar de explicação do comportamento humano, e tem de ser considerado um departamento organizado do senso comum, que pode muito bem ser chamado de uma teoria. Um modo de examinar a relação entre pensamento e linguagem é inspecionando a teoria implícita nesse tipo de explicação.

Parte da teoria lida com a explicação teleológica da ação. Perguntamo-nos por que um homem ergue seu braço; uma explicação poderia ser que ele queria atrair a atenção de um amigo. Essa explicação falharia se quem ergueu o braço não acreditasse que, erguendo seu braço, atrairia a atenção de seu amigo. Assim, a explicação completa de seu erguer o braço, ou, de qualquer modo, uma explicação mais completa, é que ele quis atrair a atenção de seu amigo e acreditou que levantando seu braço atrairia a atenção de seu amigo. Explicações desse tipo familiar têm algumas características que vale a pena enfatizar. Elas explicam o que é relativamente evidente – erguer o braço – por recurso a fatores que são, de longe, mais problemáticos: desejos e crenças. Mas, se fôssemos perguntar pelos indícios da correção das explicações, esses indícios consistiriam, no fim das contas, em mais dados a respeito do tipo de evento que está sendo explicado, a saber, outros comportamentos que são explicados pelas crenças e desejos postulados. Aludir a crenças e desejos para explicar ações é, portanto, um modo de enquadrar uma ação em um padrão de comportamento tornado coerente pela teoria. Isso certamente não significa que crenças não são nada além de padrões de comportamento, ou que os padrões relevantes possam ser definidos sem usar os conceitos de crença e desejo. Apesar disso, há um sentido claro no qual as atribuições de crença e desejo e, portanto, explicações teleológicas de crença e desejo, são supervenientes ao comportamento descrito mais amplamente.

Uma característica de explicações teleológicas não compartilhadas por explicações em geral é o modo como apelam para o conceito de *razão*. A crença e o desejo que explicam uma ação têm de ser tais que qualquer um que tivesse tal crença e desejo teria uma razão para agir daquele modo. Além disso, na explicação teleológica as descrições que fornecemos do desejo e da crença têm de exibir a racionalidade da ação, à luz do conteúdo da crença e do objeto do desejo.

A cogência de uma explicação teleológica repousa, como observado, em sua habilidade de descobrir um padrão coerente no comportamento de um agente. Coerência aqui inclui a ideia de racionalidade, tanto no sentido de que a ação a ser explicada deve ser razoável à luz dos desejos e crenças atribuídos, quanto no sentido de que os desejos e crenças atribuídos devem adequar-se um ao outro. A pressuposição metodológica da racionalidade não torna impossível atribuir

² W. Sellars, *Conceptual Change* (in: G. Pearce e P. Maynard (eds.), *Conceptual Change*. Dordrecht: D. Reidel, 1973), p. 82.

pensamentos e ações irracionais a um agente, mas impõe um preço a essas atribuições. Enfraquecemos a inteligibilidade de atribuições de quaisquer tipos de pensamentos à medida que falhamos em descobrir um padrão consistente de crenças e, finalmente, de ações, pois é apenas contra um pano de fundo desse padrão que podemos identificar pensamentos. Se virmos um homem puxando em ambas as extremidades de um pedaço de corda, podemos decidir que ele está lutando contra si mesmo, que quer mover a corda em direções incompatíveis. Essa explicação exigiria uma sustentação elaborada. Nenhum problema surge se a explicação é que ele quer arrebentar a corda.

Do ponto de vista de alguém dando explicações teleológicas das ações de outrem, claramente não faz sentido atribuir prioridade, nem a desejos, nem a crenças. Ambos são essenciais à explicação do comportamento, e nenhum está mais diretamente aberto à observação pública que o outro. Isso cria um problema, pois significa que o comportamento, que é a base evidencial para atribuições de crença e desejo, é considerado o resultado de duas forças menos abertas à observação pública. Assim, onde uma constelação de crenças e desejos racionaliza uma ação, é sempre possível encontrar uma constelação bem diferente que a racionalizará igualmente bem. Mesmo uma amostra generosa de ações ameaça deixar em aberto um número inaceitavelmente grande de explicações alternativas.

Felizmente há disponível uma teoria mais refinada, que ainda está firmemente baseada no senso comum: a teoria da preferência ou da tomada de decisões sob incerteza. A teoria foi tornada precisa pela primeira vez por Frank Ramsey, embora ele a concebesse como uma questão de fornecer uma fundamentação para o conceito de probabilidade, antes que como uma parte da psicologia filosófica.³ A teoria de Ramsey opera quantificando força de preferência e grau de crença, de modo a dar sentido à ideia natural de que, ao escolhermos um curso de ação, consideramos não apenas quão desejáveis vários resultados são, mas, também, quão aptos os cursos de ação disponíveis são em produzir esses resultados. A teoria não supõe que possamos julgar diretamente graus de crença ou comparar valores numericamente. Antes, postula um padrão razoável de preferências entre cursos de ação, e mostra como construir um sistema de crenças e desejos quantificados para explicar as escolhas. Dadas as condições idealizadas postuladas pela teoria, o método de Ramsey torna possível identificar univocamente as crenças e desejos relevantes. Em vez de falar de postulações, poderíamos colocar a questão deste modo: na medida em que podemos ver as ações de um agente como pertencendo a um padrão consistente (racional) de um certo tipo, podemos explicar essas ações em termos de um sistema de crenças e desejos quantificados.

Voltaremos mais adiante à teoria da decisão; agora é hora de voltar à questão de como a fala é interpretada. A meta imediata de uma teoria da interpretação é fornecer o significado de um proferimento arbitrário por um membro de uma comunidade linguística. Central à interpretação, eu argumentei, é uma teoria da verdade que satisfaça a Convenção T de Tarski (modificada de certas maneiras de modo a aplicar-se à linguagem natural). Pode-se tomar uma teoria dessas como fornecendo uma interpretação a cada frase que um falante possa proferir.^b Para

³ F. P. Ramsey, *Truth and Probability* (reimpresso em *Foundations of Mathematics*. New York: Humanities Press, 1950).

^b Na versão original de 1975 esta passagem está redigida assim: “Uma teoria dessas produz, para cada proferimento, de cada frase da linguagem, um teorema da forma: ‘Um proferimento da frase *s* por um falante *x* no tempo *t* é verdadeira se, e somente se, —’. Aqui ‘*s*’ deve ser substituída por uma

pertencer a uma comunidade de falantes – para ser um intérprete da fala de outros – é necessário, com efeito, conhecer tal teoria, e saber que se trata de uma teoria do tipo correto.⁴

Uma teoria da interpretação, semelhantemente a uma teoria da ação, permite-nos redescrever certos eventos de um modo revelador. Assim como uma teoria da ação pode responder à questão de o que um agente estava fazendo quando ergueu seu braço, redescrevendo o ato como uma tentativa de chamar a atenção de seu amigo, assim também um método de interpretação pode nos levar a redescrever o proferimento de certos sons como um ato de dizer que a neve é branca. Nesse ponto, entretanto, a analogia falha. Pois a teoria da decisão também pode explicar ações, embora não seja de modo algum claro como uma teoria da interpretação pode explicar as palavras “A neve é branca”, proferidas por um falante. Mas, afinal, isso era de se esperar, pois proferir palavras é uma ação e, como tal, precisa buscar sua explicação teleológica em crenças e desejos. A interpretação não é irrelevante para a explicação teleológica da fala, uma vez que, para explicar por que alguém disse algo, precisamos saber, entre outras coisas, sua própria interpretação do que disse, isto é, o que acredita que suas palavras significam nas circunstâncias em que as fala. Naturalmente isso envolverá algumas de suas crenças sobre como outros interpretarão suas palavras.

A interdependência da teoria da ação com a interpretação emergirá de outra maneira se nos perguntarmos como um método de interpretação é testado. Ao final, a resposta deve ser que ele ajuda a ordenar nossa compreensão do comportamento. Mas num estágio intermediário, podemos ver que a atitude de *sustentar como verdadeira*, ou *aceitar como verdadeira*, dirigida a frases, deve desempenhar um papel central em dar forma a uma teoria. Por um lado, a maioria dos usos da linguagem nos diz diretamente, ou ilumina a questão sobre se um falante julga verdadeira uma frase. Se o propósito de um falante é dar informação, ou fazer uma asserção honesta, então ele normalmente acredita que está proferindo uma frase verdadeira naquelas circunstâncias. Se profere uma ordem, geralmente podemos tomar isso como mostrando que julga ser falsa uma frase (intimamente relacionada à frase proferida); semelhantemente para muitos casos de engodo. Quando uma pergunta é feita, ela geralmente indica que quem pergunta não sabe se uma certa frase é verdadeira, e assim por diante. A fim de inferir a partir de tais indícios que um falante julga verdadeira uma frase, precisamos saber muito acerca de seus desejos e crenças, mas não precisamos saber o que suas palavras significam.

Por outro lado, o conhecimento das circunstâncias nas quais alguém julga frases verdadeiras é central à interpretação. Vimos no caso dos pensamentos que, embora a maioria deles não sejam crenças, é o padrão destas que nos permite identificar qualquer pensamento; de modo análogo, no caso da linguagem, embora a maioria dos proferimentos não diga respeito à verdade, é o padrão de frases julgadas verdadeiras que lhes dá o seu significado.

descrição de uma frase, e o espaço em branco por uma declaração das condições sob as quais um proferimento da frase é verdadeiro, relativamente aos parâmetros do falante e do tempo. Com vistas a interpretar um proferimento particular, não é nem necessário, nem suficiente, conhecer por inteiro a teoria: é suficiente saber o que a teoria diz serem as condições de verdade do proferimento, e saber que essas condições são implicadas por uma teoria do tipo exigido.”

⁴ Ver: Radical interpretation e Belief and the basis of meaning (ensaios 9 e 10 de *Inquiries into truth and interpretation*).

A atitude de julgar uma frase verdadeira (sob condições especificadas) relaciona crença e interpretação de um modo fundamental. Podemos saber que um falante julga uma frase verdadeira sem saber o que ele quer dizer com ela ou que crença ela expressa para ele. Mas se sabemos que ele julga verdadeira a frase e sabemos como interpretá-la, então podemos fazer uma atribuição de crença correta. Simetricamente, se sabemos qual crença uma frase julgada verdadeira expressa, sabemos como interpretá-la. O problema metodológico da interpretação é ver como encontrar, dadas as frases que um homem aceita como verdadeiras sob dadas circunstâncias, quais são suas crenças e o que suas palavras significam. A situação é novamente semelhante à situação na teoria da decisão, onde, dadas as preferências de um homem entre cursos alternativos de ação, podemos discernir suas crenças e desejos. Naturalmente, não se deve pensar que uma teoria da interpretação pode sustentar-se isoladamente, pois, como ressaltamos, não há como dizer quando uma frase é julgada verdadeira, sem ser capaz de atribuir desejos e de descrever ações como tendo intenções complexas. Essa observação não priva de interesse a teoria da interpretação, mas lhe atribui um lugar numa teoria mais abrangente da ação e do pensamento.⁵

Ainda não está claro se a interpretação é exigida para uma teoria da ação, que é a questão que nos propusemos a responder. O que é certo é que todas as maneiras convencionais de testar teorias da decisão ou da preferência sob incerteza fiam-se no uso da linguagem. É relativamente simples eliminar a necessidade de respostas verbais por parte do sujeito: pode-se considerar que ele expressou uma preferência agindo, movendo-se diretamente ao seu fim, em vez de dizer o que queria. Mas isso não pode decidir a questão sobre o que ele escolheu. Um homem que pega uma maçã em vez de uma pera, quando ambas lhe são oferecidas, pode estar expressando uma preferência pelo que está na sua esquerda antes que pelo que está na sua direita, pelo que é vermelho em vez do amarelo, pelo que é visto primeiro ou julgado mais caro. Testes repetidos podem tornar algumas leituras de suas ações mais plausíveis que outras, mas persistirá o problema de como determinar quando ele julga que dois objetos de escolha são idênticos. Testes que envolvem eventos incertos – escolhas entre apostas – são ainda mais difíceis de apresentar sem usar palavras. O psicólogo, cético de sua capacidade de ter certeza sobre como um sujeito está interpretando suas instruções, tem de acrescentar uma teoria da interpretação verbal à teoria a ser testada. Se pensarmos de todas as escolhas como revelando uma preferência de que uma frase e não outra, seja verdadeira, a teoria total resultante forneceria uma interpretação de frases e, ao mesmo tempo, atribuiria crenças e desejos, estes últimos concebidos como relacionando o agente a frases e proferimentos. Essa teoria complexa explicaria todo comportamento, verbal ou não.

Tudo isso sugere fortemente que a atribuição de desejos e crenças (e outros pensamentos) tem de ir de mãos dadas com a interpretação da fala, que nem a teoria da decisão, nem a da interpretação, pode ser desenvolvida com sucesso sem a outra. Mas resta dizer, em detalhes mais convincentes, por que a atribuição de pensamentos depende da interpretação da fala. A razão geral, e não muito informativa, é que sem a fala não podemos fazer distinções refinadas entre

⁵ A interdependência da teoria da decisão e da interpretação radical é explorada também no ensaio 10 de *Inquiries into truth and interpretation*, no ensaio 12 de *Essays on actions and events* (Oxford: Clarendon Press, 1980), e em *Toward a unified theory of meaning and action* (in: *Grazer Philosophische Studien*, 2 (1980): 1-12).

pensamentos, as quais são essenciais para as explicações que, às vezes com convicção, podemos fornecer. Nossa maneira de atribuir atitudes assegura que todo o poder expressivo da linguagem pode ser usado para fazer essas distinções. Pode-se acreditar que Scott não é o autor de *Waverley*, sem duvidar que Scott é Scott; pode-se querer ser o descobridor de uma criatura com um coração, sem se querer ser o descobridor de uma criatura com um rim. Pode-se intencionalmente morder a maçã que se tem na mão, sem intencionalmente morder a única maçã que tem uma larva em seu interior, e assim por diante. É muito difícil dar importância à intencionalidade, que é tão fecunda na atribuição de pensamentos, quando a fala não está presente. Dizemos que o cão sabe que seu dono está em casa. Mas será que ele sabe que Sr. Smith (que é seu dono), ou que o presidente do banco (que é esse mesmo dono), está em casa? Não temos nenhuma ideia real de como decidir ou dar sentido a essas questões. É muito mais difícil dizer, quando a fala não está presente, como distinguir pensamentos universais de conjunções de pensamentos, ou como atribuir pensamentos condicionais, ou pensamentos com, por assim dizer, quantificações mistas (“Ele espera que todos sejam amados por alguém”).

Essas considerações serão, provavelmente, menos persuasivas para amantes de cães que para outros, mas de qualquer modo elas não constituem um argumento. Na melhor das hipóteses, o que mostramos, ou alegamos, é que, a menos que haja comportamento que possa ser interpretado como fala, os indícios não serão adequados para justificar as distinções refinadas que estamos acostumados a fazer nas atribuições de pensamentos. Se persistirmos atribuindo desejos, crenças ou outras atitudes sob essas condições, nossas atribuições, e consequentes explicações de ações, serão seriamente subdeterminadas, de modo que muitos sistemas alternativos de explicação, muitas explicações alternativas, estarão igualmente justificados pelos dados disponíveis. Talvez isso seja tudo o que possamos dizer contra a atribuição de pensamentos a criaturas mudas, mas eu não penso assim.

Antes de avançar, quero considerar uma objeção possível à linha geral de argumentação que estive perseguindo. Suponha que concedamos, o objetor diz, que comportamentos muito complexos, não observados em crianças pequenas e elefantes, é necessário se quisermos encontrar uma aplicação para o aparato completo disponível para a atribuição de pensamentos. Ainda assim, pode-se dizer que o esboço de como funciona a interpretação não mostra que essa complexidade tenha de ser vista como conectada à linguagem. A razão é que o esboço faz com que coisas demais dependam da atitude especial de ser pensado como verdadeiro. O indício mais direto da existência dessa atitude é a asserção honesta. Mas, então, pareceria que poderíamos tratar como fala o comportamento de criaturas que nunca fizeram nada com a linguagem, exceto asserções honestas. Alguns filósofos sonham com essas tristes tribos, mas estaríamos corretos em dizer que teriam uma linguagem? O que se perdeu de vista é o que pode ser chamado de *a autonomia do significado*. Uma vez que uma frase é entendida, pode-se proferi-la com praticamente qualquer propósito extralinguístico. Um instrumento, ao qual se pudesse dar somente um uso, careceria de autonomia do significado. Isso equivale a dizer que não deveria ser contado como uma linguagem. Assim, a complexidade de comportamento necessário para dar um escopo completo a atribuições de pensamento não precisa ter, no fim das contas, exatamente a mesma complexidade que permita, ou exija, interpretação como linguagem.

Concordo com o objetor hipotético que a autonomia do significado é essencial à linguagem; de fato, é isso que explica, em grande medida, por que o significado linguístico não pode ser definido ou analisado com base em intenções e crenças extralinguísticas. Mas o objetor falha em distinguir entre uma linguagem que *poderia* ser usada com apenas um propósito, e uma que *é* usada com apenas um propósito. Um instrumento que pudesse ser usado para apenas um propósito não seria uma linguagem. Mas a asserção honesta poderia por só render uma teoria da interpretação e, assim, também uma linguagem, da qual, embora capaz de mais, nunca fossem feitos outros usos. (Como uma questão prática, o evento é impensável. Alguém que sabe sob que condições suas frases são socialmente verdadeiras, não pode deixar de apreender e valer-se de possibilidades de asserção desonesta – ou de fazer piadas, contar histórias, provocar, exagerar, insultar e todo o restante dessas coisas divertidas.)

Um método de interpretação nos diz que, para falantes do português, um proferimento de “Está chovendo”, por um falante x no tempo t , é verdadeiro se, e somente se, está chovendo (próximo de x) em t . Estar munido dessa informação, e saber que outros também sabem, é saber o que um proferimento significa independentemente de saber os propósitos que o incitaram. A autonomia do significado também ajuda a explicar como é possível, pelo uso da linguagem, atribuir pensamentos. Suponha que alguém profere assertivamente a frase “A neve é branca”. Sabendo as condições sob as quais tal proferimento é verdadeiro, posso acrescentar, se eu quiser, “Eu também acredito nisso”, atribuindo, assim, uma crença a mim mesmo. Nesse caso, ambos podemos ter asserido que a neve é branca, mas não necessariamente com a mesma força de autoatribuição. O outro pode dizer zombeteiramente, expressando descrença, “A neve é branca” – e posso novamente atribuir uma crença a mim mesmo dizendo “Mas *eu* acredito nisso”. Isso pode funcionar bem de um outro modo: se posso aproveitar um proferimento de outra pessoa para atribuir uma crença a mim mesmo, posso usar um proferimento de minha autoria para atribuir uma crença a mais alguém. Primeiro profiro uma frase, talvez “A neve é branca”, e, então, acrescento “Ele acredita nisso”. O primeiro proferimento pode ser ou não uma asserção; de qualquer modo, ele não atribui uma crença a ninguém (embora, caso seja uma asserção, então eu *represento* a mim mesmo como acreditando que a neve é branca). Mas se meu comentário “Ele acredita nisso” é uma asserção, eu atribuí uma crença a alguém. Por fim, não há nenhum impedimento em eu atribuir uma crença a mim mesmo, dizendo primeiro “A neve é branca”, e depois acrescentando “Eu acredito nisso”.

Em todos esses exemplos, tomo a palavra “nisso” [*that*] como se referindo demonstrativamente a um proferimento, quer se trate de um proferimento feito por quem disse “nisso”, quer por outro falante. O “nisso” não pode referir-se a uma frase, porque, como Church salientou em casos similares, a referência teria de ser então relativizada a uma linguagem, uma vez que uma frase pode ter diferentes significados em diferentes linguagens,⁶ e também, e mais obviamente, porque a mesma frase pode ter diferentes valores de verdade na mesma linguagem.

O que a referência demonstrativa a proferimentos faz nesse tipo de caso recém-considerado, ela também pode fazer quando a estrutura superficial é alterada

⁶ A. Church, On Carnap's Analysis of Statements of Assertion and Belief, *Analysis*, 10 (1950): 97-99.

para algo como “Eu acredito que [*that*] a neve é branca”^c ou “Ele acredita que [*that*] a neve é branca”. Penso que também nesses exemplos devemos ver o “que” [*that*] como um demonstrativo, agora se referindo adiante, a um proferimento que está prestes a ser produzido. Assim, a forma lógica *standard* das atribuições de atitudes é aquela de dois proferimentos unidos parataticamente. Não há conetivo, embora o primeiro proferimento contenha uma referência ao segundo. (Observações similares valem, certamente, para registros escritos de frases).

Discuti essa análise das atribuições verbais de atitudes em outro lugar, e não há necessidade de repetir aqui os argumentos e explicações.⁷ Trata-se de uma análise com suas próprias dificuldades, especialmente quando se analisa quantificação na frase contida, mas penso que essas dificuldades podem ser superadas, enquanto são preservadas os aspectos atraentes da ideia. Quero enfatizar aqui o ponto que conecta a análise paratática de atribuições de atitudes com o tema presente. A análise proposta relaciona diretamente a característica de autonomia do significado com nossa habilidade de descrever e atribuir pensamentos, uma vez que é somente porque a interpretação de uma frase é independente de seu uso, que o proferimento de uma frase pode ter serventia na descrição a atitude de outros. Se minha análise está correta, podemos dispensar a concepção improvável (mas comum) segundo a qual uma frase encerrada em uma cláusula-“que” necessita uma interpretação inteiramente diferente daquela que funciona para ela outros contextos. Como em contextos ordinários frases não são nomes ou descrições, podemos, em particular, rejeitar a suposição de que atitudes têm objetos tais como proposições, as quais, podem sustentar cláusulas -“que” nomeiam ou descrevem. Não deveria haver a tentação de chamar o proferimento a que se faz referência, de acordo com a análise paratática, de objeto da atitude atribuída.

Aqui uma solução fácil para o nosso problema sobre a relação entre pensamentos e fala sugere-se a si mesma. Um modo de ver a análise paratática, um modo proposto por Quine em *Palavra e objeto*, é este: quando um falante atribui uma atitude a uma pessoa, o que ele faz é arremedar ou imitar um ato de fala real ou possível dessa pessoa.⁸ O discurso indireto é o melhor exemplo, e a asserção é outro bom exemplo. Suponha que eu diga “Heródoto asseriu que o Nilo nasce nas Montanhas da Lua”. Meu segundo proferimento – meu proferimento recém-feito de “O Nilo nasce nas Montanhas da Lua” – tem de, se minha atribuição a Heródoto está correta, guardar uma certa relação com um proferimento de Heródoto: tem de ser, em algum sentido apropriado, uma tradução dele. Uma vez que, supondo-se que a atribuição está correta, Heródoto e eu *dissemos o mesmo* [*are samesayers*], meu proferimento imitou o dele. Não com respeito à força, certamente, uma vez que não asseri nada a respeito do Nilo. A identidade diz respeito ao conteúdo de nossos proferimentos. Se voltamos a outras atitudes, a situação é mais complicada, pois tipicamente não há um proferimento a imitar. Se afirmo “Jones acredita que a neve é branca”, meu proferimento de “A neve é branca” pode não ter um proferimento real de Jones para imitar. Ainda assim, poderíamos seguir a linha de raciocínio segundo a qual o que afirmo é que Jones estaria dizendo honestamente o que pensa, se ele proferisse uma frase que traduzisse a minha. Dados alguns pressupostos delicados

^c Em inglês, a palavra para “nisso” e “que” nos exemplos de Davidson é a mesma, *that*: “I believe *that*” (Eu acredito *nisso*) e “I believe *that* snow is white” (Eu acredito *que* a neve é branca).

⁷ Ver On saying that (ensaio 7 de *Inquiries into truth and interpretation*).

⁸ W. V. Quine, *Word and object* (Cambridge, Mass.: Harvard University Press, 1960), p. 219.

sobre as condições em que um condicional subjuntivo desses é verdadeiro, poderíamos concluir que apenas alguém com uma linguagem poderia ter um pensamento, uma vez que ter um pensamento seria ter uma disposição para proferir certas frases sob dadas circunstâncias com a força apropriada.

Poderíamos seguir essa linha de raciocínio, mas infelizmente não parece haver uma razão clara para fazê-lo. Propusemos-nos a encontrar um argumento para mostrar que apenas criaturas com fala têm pensamentos. O que acaba de ser esboçado não é um argumento, mas uma proposta, e uma proposta que não precisamos aceitar. A análise paratática da forma lógica de atribuições de atitudes pode seguir em frente sem a teoria de proferimentos como imitação [*mimic-theory of utterance*]. Quando digo “Jones acredita que a neve é branca”, descrevo o estado mental de Jones diretamente: é de fato o estado mental em que alguém está se pudesse honestamente asserir “A neve é branca”, caso falasse português, mas esse também pode ser o estado em que uma criatura sem linguagem poderia estar.

A fim de apresentar meu argumento [*point*] final principal, preciso retornar a um aspecto da interpretação até aqui negligenciado. Observei que a atitude de julgar verdadeira, dirigida a frases sob circunstâncias específicas, é a base para a interpretação, mas não disse como ela pode exercer essa função. A dificuldade, recordando, é que uma frase é julgada verdadeira devido a dois fatores: aquilo que quem julga toma a frase como significando, e o que ele acredita. De modo a esclarecer as coisas, o que é necessário é um método para manter fixo um fator enquanto o outro é estudado.

O pertencimento a uma comunidade linguística depende da capacidade de interpretar os proferimentos dos membros do grupo, e um método torna-se disponível se alguém tem, e sabe que tem uma teoria que fornece as condições de verdade, mais ou menos ao estilo de Tarski, para todas as frases (relativizadas, como sempre, ao tempo e ao falante). A teoria é desde que ela acarrete, por meios finitamente declarados, teoremas da forma familiar: “‘Está chovendo’ é verdadeira para um falante x em t se, e somente se, está chovendo (próximo a x) em t ”. A base evidencial para uma tal teoria diz respeito a frases julgadas verdadeiras, fatos como o seguinte: “‘Está chovendo’ foi julgada verdadeira por Smith às 8h de 26 de agosto e chovia próximo a Smith naquele momento”. Seria possível gerar uma teoria correta simplesmente considerando-se frases como sendo verdadeiras quando são julgadas verdadeiras, desde que (1) houvesse uma teoria que satisfizesse as restrições formais e fosse consistente, desse modo, com a evidência, e (2) todos os falantes julgassem uma frase verdadeira somente quando tal frase fosse verdadeira – desde que, isto é, todas as crenças, ao menos à medida que possam ser expressas, fossem corretas.

Mas certamente não se pode supor que os falantes jamais têm crenças falsas. O erro é o que dá à crença sua razão de ser [*point*]. Podemos, entretanto, tomar como dado que a *maioria* das crenças são corretas. A razão para isso é que uma crença é identificada por seu lugar em um padrão de crenças; é esse padrão que determina o objeto da crença, aquilo do que a crença trata. Antes que um objeto ou aspecto do mundo possa tornar-se parte do conteúdo de uma crença (verdadeira ou falsa), tem de haver infundáveis crenças verdadeiras sobre o mesmo. Crenças falsas tendem a minar a identificação do conteúdo; minar, portanto, a validade de uma descrição da crença como sendo a respeito daquele assunto. E assim, por sua vez, crenças falsas minam a alegação de que uma crença conectada a ela é falsa. Por exemplo, que clareza temos de que os antigos – alguns antigos – acreditaram

que a terra era plana? *Esta terra?* Bem, esta nossa terra é parte do sistema solar, um sistema parcialmente identificado pelo fato de que é uma multidão de grandes corpos, frios e sólidos, circulando ao redor de uma estrela quente e enorme. Se alguém não acredita em *nada* disso sobre a terra, é certo que é a respeito da terra que está pensando? Uma resposta não é necessária. Alcançamos nosso objetivo se esse tipo de consideração sobre crenças relacionadas pode abalar a confiança que se tem de que os antigos acreditavam que a terra era plana. Não se trata de que uma crença falsa qualquer necessariamente destrói nossa capacidade de identificar as demais crenças, mas que a inteligibilidade dessas identificações tem de depender de um pano de fundo de crenças verdadeiras, em sua maioria não-mencionadas e não-questionadas. Em outras palavras: quanto maior o número de coisas sobre as quais um crente está certo, mais nítidos serão seus erros. Erros em excesso simplesmente obscurecem o foco.

O que torna a interpretação possível é, então, o fato de que podemos rejeitar *a priori* a chance de erro massivo. Uma teoria da interpretação não pode estar correta se faz com que alguém assinta a um número muito grande frases falsas: tem de geralmente ser o caso que uma frase é verdadeira quando um falante a julga verdadeira. Pelo que vale [*so far as it goes*], conta a favor de um método de interpretação que ele apenas tome uma frase como verdadeira quando os falantes a julgam verdadeira. Mas certamente o falante pode estar errado; e assim também o intérprete. Assim, ao final, o que deve contar em favor de um método de interpretação é que ele coloca em geral o intérprete em acordo com o falante: segundo o método, o falante julga uma frase verdadeira sob condições especificadas, e essas condições se realizam, na opinião do intérprete, apenas quando o falante julga que a frase é verdadeira.

Nenhuma teoria simples pode colocar um falante e um intérprete em perfeito acordo, e, assim, uma teoria viável tem de supor, de tempos em tempos, o erro por parte de um ou outro. O preceito metodológico básico é, portanto, que uma boa teoria da interpretação maximiza o acordo. Ou, dado que as frases são infinitas em número, e dadas considerações adicionais a serem feitas, *otimizar* talvez seja uma palavra melhor.

Alguns desacordos são mais destrutivos ao entendimento que outros, e uma teoria sofisticada naturalmente tem de levar isso em consideração. Desacordo sobre questões teóricas pode (em alguns casos) ser mais tolerável que desacordo sobre o que é mais evidente; desacordo sobre como as coisas parecem ou afiguram-se é menos tolerável do que um desacordo sobre como elas são; desacordo sobre a verdade de atribuições de certas atitudes a um falante por esse mesmo falante pode não ser de modo algum tolerável, ou ser quase intolerável. É impossível simplificar as considerações que são relevantes, pois tudo o que sabemos ou acreditamos sobre o modo como os indícios sustentam crenças pode ser usado para decidir onde é melhor que a teoria permita erros, e quais erros são menos destrutivos do entendimento. A metodologia da interpretação é, a esse respeito, nada mais que a epistemologia vista no espelho do significado.

O intérprete que supõe que seu método pode ser posto a serviço de uma comunidade linguística batalhará por uma teoria que otimize o acordo por toda a comunidade. Uma vez que a comunicação fácil tem valor de sobrevivência, ele pode esperar que os usos no interior de uma comunidade favoreçam teorias da interpretação simples e comuns.

Se essa abordagem da interpretação radical está correta, ao menos em linhas gerais, então temos de reconhecer que os conceitos de verdade objetiva e de erro emergem necessariamente no contexto da interpretação. A distinção entre uma frase ser julgada verdadeira e ser, de fato, verdadeira é essencial à existência de um sistema interpessoal de comunicação, e quando, em casos individuais, há uma diferença, esta tem de ser contada como erro. Uma vez que a atitude de julgar verdadeira é a mesma, quer a frase seja verdadeira ou não, ela corresponde diretamente à crença. O conceito de crença, assim, permanece pronto para ajustar [*take up the slack*] a verdade objetiva ao tomar por verdadeiro, e apenas chegamos a entendê-lo nessa conexão.

Obtemos a ideia de crença somente a partir do papel da crença na interpretação da linguagem, pois como atitude privada ela não é inteligível, exceto como um ajuste à norma pública fornecida pela linguagem. Segue-se que uma criatura tem de pertencer a uma comunidade linguística de modo a ter o conceito e crença. E dada a dependência de outras atitudes sobre a crença, podemos dizer, de modo mais geral, que somente uma criatura que pode interpretar a fala pode ter o conceito de um pensamento.

Pode uma criatura ter uma crença se ela não tem o conceito de crença? Parece-me que não, e por esta razão. Alguém não pode ter uma crença a menos que entenda a possibilidade de estar enganado, e isso exige apreender o contraste entre verdade e erro – crença verdadeira e crença falsa. Mas esse contraste, argumentei, somente pode emergir no contexto da interpretação, o qual, sozinho, nos força para a ideia de uma verdade objetiva, pública.

É frequentemente pensado, erroneamente, que o conceito de verdade é redundante, que não há diferença entre asserir que uma frase *s* é verdadeira, e usar *s* para fazer uma asserção. O que pode estar certo é uma teoria da redundância da crença, que acreditar que *p* não deve ser distinguido da crença de que *p* é verdadeira. Essa noção de verdade não é a noção semântica: a linguagem não está diretamente na pintura. Mas ela está apenas ligeiramente [*only just out*] fora da pintura; ela é parte da moldura. Pois a noção de uma crença verdadeira depende da noção de uma frase verdadeira, e isso, por sua vez, não pode existir sem uma linguagem compartilhada. Como o Ulysses de Shakespeare diz: “[...] nenhum homem é senhor de nada, ainda que muito seja ou tenha, até que comunique suas coisas a outrem; nem mesmo ele as conhece de todo, até que as veja tomar forma no aplauso em que se estendem”.⁹

⁹ Shakespeare, W., *Troilus and Cressida*, III, iii.