

# La mente mecánica

*Introducción filosófica a mentes,  
máquinas y representación mental*

por TIM CRANE



FONDO DE CULTURA ECONÓMICA

## I. EL ROMPECABEZAS DE LA REPRESENTACIÓN

CUANDO LA NASA envió el *Pioneer 10* al espacio para explorar el sistema solar en 1972, se puso a bordo una placa de metal grabada con varias imágenes y signos. Parte de la placa era un diagrama de un átomo de hidrógeno, en tanto que otra era un diagrama de los tamaños relativos de los planetas en nuestro sistema solar, indicando el planeta del cual procedía el *Pioneer 10*. La imagen mayor en esta placa era la silueta de un hombre y una mujer desnudos, con el hombre alzando la mano para saludar. La idea que había detrás de esto era que cuando el *Pioneer 10* acabase por abandonar el sistema solar, podría continuar un camino cualquiera a través del espacio, tal vez para ser descubierto en millones de años por alguna forma vital ajena. Y tal vez estos alienígenas serían inteligentes y conseguirían comprender los diagramas, reconocer el grado de nuestro conocimiento científico y darse cuenta de que nuestras intenciones hacia ellos, quienes fueran, eran pacíficas.

Me parece que hay algo muy humorístico en esta historia. Supóngase que el *Pioneer 10* alcanzase alguna estrella distante. Y supóngase que la estrella tuviera un planeta con condiciones susceptibles de sostener la vida. Y finalmente supóngase que alguna forma de vida de este planeta fuera inteligente y tuviera alguna especie de órganos sensoriales con los cuales pudiera percibir la lámina volante. Todo esto es sumamente improbable. Pero incluso al admitir estas

improbables suposiciones, ¿no parecería aún *más* imposible que los alienígenas consiguieran comprender qué significan los símbolos de la placa?

Piénsese en algunas de las cosas que deberían comprender. Tendrían que comprender que los símbolos sobre la placa *eran* símbolos, que aspiraban a representar cosas y no eran sencillamente raspaduras en la placa, o simple decorado. Una vez que los alienígenas supieran que eran símbolos, tendrían que comprender de qué clase de símbolos se trataba: por ejemplo, que el diagrama del átomo de hidrógeno era un diagrama científico y no una simple estampa. Entonces deberían tener alguna idea de qué clases de cosas representaban los símbolos: que el dibujo del hombre y la mujer simbolizaban formas vitales en lugar de elementos químicos, que el diagrama del sistema solar simbolizaba nuestra parte del universo antes que la forma de quienes diseñaron el vehículo espacial. Y —quizá lo más absurdo de todo—, aun cuando se hiciesen una idea de lo que el hombre y la mujer eran, tendrían que reconocer que la mano alzada era un signo de saludo pacífico y no de impaciencia, de agresión o de desdén, o sencillamente que era la posición normal de esta parte del cuerpo.

Cuando se considera todo esto, ¿no parece aún más improbable que los alienígenas comprendieran los símbolos que la llegada del vehículo espacial a un planeta con vida inteligente, para empezar?

Una cosa que ilustra esta historia, en mi opinión, es algo acerca del problema filosófico o el rompecabezas de la representación. Los dibujos y símbolos de la placa representan cosas —átomos, seres humanos, el sistema solar—, pero la historia sugiere que hay algo desconcertante acerca del modo como lo logran. Porque cuando nos imaginamos en

la posición de los alienígenas, nos damos cuenta de que no podemos decir qué representan estos símbolos con sólo mirarlos. Ningún grado de escrutinio de las marcas en la placa puede revelar que estas marcas representan a un hombre y aquéllas representan a una mujer, y que estas otras marcas representan un átomo de hidrógeno. Las marcas de la placa pueden ser entendidas de muchas maneras, pero parece que nada en las marcas *mismas* revela cómo entenderlas. Ludwig Wittgenstein, cuya filosofía estaba dominada por cuestiones acerca de la representación, expresó brevemente: “Cada signo *por sí mismo* parece muerto; ¿qué le da vida?”<sup>1</sup>

El rompecabezas filosófico acerca de la representación puede ser planteado sencillamente preguntando cómo es posible que una cosa represente otra. Planteada así, la cuestión puede parecer un poco oscura, y puede ser difícil ver exactamente qué es desconcertante acerca de ella. Una razón para esto es que la representación sea un hecho tan familiar en nuestras vidas. Palabras dichas y escritas, imágenes, símbolos, gestos, expresiones faciales, pueden todas verse como representaciones a partir del tejido de nuestra vida cotidiana. Sólo cuando iniciamos la reflexión acerca de cosas como la historia del *Pioneer 10* empezamos a ver cuán desconcertantes son realmente. Nuestras palabras, imágenes, expresiones, y así por el estilo, representan, significan o revelan cosas, pero ¿cómo?

Por un lado, la representación acude naturalmente a nosotros. Cuando hablamos uno con otro, o miramos una imagen, lo que representa suele ser inmediato, y no algo que tengamos que averiguar. Por otra parte, palabras e imá-

<sup>1</sup> Wittgenstein, *Philosophical Investigations* (Oxford, Blackwell, 1953), § 432.

genes no son sino pautas físicas: vibraciones en el aire, marcas en papel, piedra, plástico, película o (como en el *Pioneer 10*) placas metálicas. Tómese el ejemplo de las palabras. Es evidente que no hay nada en las pautas físicas de las palabras mismas que las hagan representar como lo hacen. Los niños a menudo se familiarizan con este hecho cuando repiten palabras para ellos mismos una y otra vez, hasta que parecen “perder” su sentido. Quienquiera que haya aprendido una lengua extranjera reconocerá que, por natural que parezca en el caso de nuestro propio lenguaje, las palabras no tienen significado *en sí y por sí mismas*. O, como dicen los filósofos, no tienen significado “intrínseco”.

Por un lado, pues, la representación parece natural, espontánea y sin problemas. Por otra parte, la representación parece innatural, enrevesada y misteriosa. Igual que con los conceptos de tiempo, verdad y existencia (por ejemplo), el concepto de representación plantea un rompecabezas característico de la filosofía: lo que parece natural y evidente en nuestras vidas se torna, al reflexionar, hondamente misterioso.

El problema filosófico de la representación es un tema fundamental de este libro. Es uno de los problemas centrales de la filosofía contemporánea de la mente. Y otros muchos puntos filosóficos se acumulan alrededor de este problema: el lugar de la mente en la naturaleza, la relación entre pensamiento y lenguaje, la naturaleza de nuestra comprensión entre uno y otro, el problema de la conciencia y la posibilidad de máquinas pensantes. Todos estos puntos serán tocados aquí. La meta de este capítulo es aguzar nuestra comprensión del problema de la representación mostrando cómo ciertas soluciones en apariencia evidentes al respecto sólo conducen a nuevos problemas.

### LA IDEA DE REPRESENTACIÓN

Comenzaré afirmando algunas cosas muy generales acerca de la idea de representación. No hay que alarmarse de enunciar lo evidente: una representación es algo que representa a otra cosa. No digo que una representación sea algo que representa algo *más*, porque una representación puede representarse a sí misma. (Hay un ejemplo famoso de la “Paradoja del Mentiroso”: “Esta oración es falsa” representa la oración misma.) El caso normal es donde una cosa —la representación misma— representa otra cosa, que podríamos llamar *objeto* de la representación. Por lo tanto podemos plantear dos cuestiones: una acerca de la naturaleza de las representaciones y una acerca de la naturaleza de los objetos de la representación.

¿Qué clase de cosas pueden ser las representaciones? Ya he mencionado las palabras y las imágenes, que son quizá los ejemplos más evidentes. Por supuesto, hay otros muchos casos. El diagrama de la placa del átomo de hidrógeno en el *Pioneer 10* no es un manojo de palabras ni una imagen, sino que representa el átomo de hidrógeno. Los numerales, como 15, 23, 1 001, etc., representan números. Los números pueden representar otras cosas también: por ejemplo, un numeral puede representar la longitud de un objeto (en metros o en pies) y un triple número puede representar un matiz particular de color representando su grado de matiz, saturación y claridad. Las estructuras de datos de una computadora pueden representar texto o números o imágenes. Los anillos de un árbol pueden representar su edad. Una bandera puede representar una nación. Una demostración política puede representar agresión. Un fragmento de música

ca puede representar un talante de insoportable melancolía. Las flores pueden representar la pena. Una ojeada o una expresión facial puede representar irritación. Y, como veremos, un estado mental —una creencia, una esperanza, un deseo o un anhelo— puede representar casi cualquier cosa.

Hay tantas clases de cosas que pueden ser representaciones, que costaría más de un libro discutir las todas. Y, por supuesto, no pretenderé hacer esto. Enfocaré ejemplos sencillos de representación en el lenguaje y en el pensamiento. Por ejemplo, hablaré acerca de cómo es que puedo utilizar una palabra para representar una persona particular o cómo puedo preguntar (digamos) acerca de un perro. Me fijaré estos sencillos ejemplos porque los problemas filosóficos acerca de la representación surgen incluso en los casos más sencillos. Introducir casos más complejos —tal como de qué manera una pieza de música puede representar una actitud— en esta etapa sólo haría el punto más difícil y confundiría más la mente que antes. Ignorar estos casos complejos no significa que considere que no son importantes o que no tienen interés.<sup>2</sup>

Ahora nuestra segunda cuestión: ¿qué clase de cosas pueden ser objetos de representación? La respuesta es, evidentemente, casi cualquier cosa. Las palabras y las imágenes pueden representar un objeto físico, como una persona o una casa. Pueden representar un rasgo o una propiedad de un objeto físico; por ejemplo, la forma de una persona o el color de una casa. Oraciones como “Alguien está en mi casa” pueden representar lo que podríamos llamar hechos, situaciones o estados de cosas: en este caso, el hecho de que

<sup>2</sup> Acerca de la cuestión, por ejemplo, de cómo puede la música expresar emoción, véase Malcolm Budd, *Music and the Emotions* (Londres, Routledge, 1986).

alguien esté en mi casa. Los objetos no físicos pueden ser representados también: si hay números, no son evidentemente objetos físicos (¿cuándo en el mundo físico se da el número 3?). Representaciones —tales como palabras, imágenes, música y expresiones faciales— pueden representar talantes, sentimientos y emociones. Y las representaciones pueden representar cosas que no existen. Puedo pensar —o sea representar— en unicornios, dragones y el mayor número primo. Ninguna de estas cosas existe; pero pueden todas ser “objetos” de representación.

Este último ejemplo indica un rasgo curioso de la representación. Frente a ella, la expresión “*X* representa *Y*” sugiere que la representación es una *relación* entre dos cosas. Sin embargo, una relación entre dos cosas normalmente implica que estas dos cosas existen. Tómese la relación de *besar*: si beso a Santa Claus, entonces Santa Claus y yo debemos existir ambos. Y el hecho de que Santa Claus no exista explica por qué no puedo besarlo.

Pero esto no es cierto acerca de la representación: si pienso en Santa Claus, y por lo tanto lo represento, no se sigue que Santa Claus exista. La no existencia de Santa Claus no es obstáculo para mi representación de él, como sí lo es para besarlo. De esta manera, la representación parece muy diferente de otras relaciones. Como veremos más adelante, muchos filósofos han tomado este aspecto de la representación por centro de su naturaleza.

Así, hay muchas clases de representaciones, y muchas clases de cosas que pueden ser objetos de representación. ¿Cómo podemos hacer cualquier progreso en la comprensión de la representación? Hay dos clases de preguntas que podemos hacer.

Primero, podemos preguntar *cómo* alguna clase de re-

presentación —imágenes, palabras o lo que sea— consigue representar. Lo que queremos conocer es qué pasa con esta clase de representación que hace que tenga su papel representativo. (Como ejemplo considero más adelante la idea de que las imágenes podrían representar cosas *pareciéndose* a ellas.) Evidentemente, no supondremos que la historia de una forma de representación se aplicará por necesidad a todas las demás formas: el modo como las imágenes representan no será el mismo que la manera como la música representa, por ejemplo.

En segundo lugar, podemos preguntar si alguna forma particular de representación es más *básica* o *fundamental* que las otras. Es decir, si podemos explicar algunas clases de representación en términos de otras. Por ejemplo, una cuestión de la filosofía contemporánea es si podemos explicar el modo como el lenguaje representa en términos de capacidades representativas de estados mentales, o si necesitamos explicar la representación mental en términos de lenguaje. Si hay una clase de representación que sea más fundamental que las otras clases, entonces estamos claramente en el camino de comprender la representación en conjunto.

Mi propio punto de vista es que la representación mental —la representación del mundo por estados mentales— es la forma más fundamental de representación. Para ver cómo éste puede ser un punto de vista razonable, necesitamos mirar brevemente una representación pictórica y lingüística.

#### IMÁGENES Y PARECIDO

Tal como se presentan las cosas, la manera como las imágenes representan parece ser más rectilínea que otras formas

de representación. Pues, en tanto que nada hay de intrínseco en la palabra *perro* que haga que represente perros, de seguro hay algo intrínseco en un cuadro de un perro que hace que parezca un perro, esto es, *aquello a lo cual se asemeja la imagen*. Las imágenes de perros se ven aproximadamente como perros, parecen perros de alguna manera, y lo hacen a causa de sus rasgos intrínsecos: su forma, su color, y así sucesivamente. Tal vez, entonces, una imagen representa lo que representa por parecerse a dicha cosa.

La idea de que una imagen representa pareciéndose sería una respuesta a la primera clase de cuestión mencionada anteriormente: ¿cómo una clase particular de representación consigue representar? La respuesta es: las imágenes representan cosas pareciéndose a dichas cosas. (Esta respuesta puede entonces usarse como fundamento para una respuesta a la segunda cuestión; la sugerencia será que todas las formas de representación pueden ser explicadas en términos de representación pictórica. Pero, como veremos más adelante, esta idea no tiene esperanza.) Llamemos a esta idea “teoría del parecido para la representación pictórica”, o “teoría del parecido”, por usar pocas palabras. Discutir la teoría del parecido de manera más precisa requiere un poco de terminología filosófica básica.

Los filósofos distinguen entre dos maneras de que la verdad de una opinión dependa de la verdad de otra. Llamamos a estas dos maneras condiciones “necesaria” y “suficiente”. Decir que una pretensión particular *A* es una condición *necesaria* para alguna otra pretensión *B*, es decir esto: *B* es verdad sólo si *A* es verdad también. Intuitivamente, *B* no será verdad sin que *A* lo sea, de suerte que la verdad de *A* es *necesaria* (es decir, necesitada, requerida) para la verdad de *B*.

Decir que *A* es una condición *suficiente* para *B* es decir

esto: si  $A$  es verdadera, entonces  $B$  es verdadera también. Intuitivamente, la verdad de  $A$  asegura la verdad de  $B$  o, en otras palabras, la verdad de  $A$  *basta* para la verdad de  $B$ . Decir que  $A$  es una condición necesaria y suficiente para la verdad de  $B$  es decir esto: si  $A$  es verdad,  $B$  es verdad, y si  $B$  es verdad,  $A$  es verdad. (Esto es expresado a veces diciendo que “ $A$  es verdad si y sólo si  $B$  es verdad” y “si y sólo si es abreviado en ocasiones como ‘sic’”.)

Ilustremos esta distinción con un ejemplo. Si estoy en Londres, entonces estoy en Inglaterra. Así, estar en Inglaterra es una *condición necesaria* para estar en Londres: no puedo precisamente estar en Londres sin estar en Inglaterra. Análogamente, estar en Londres es una *condición suficiente* para estar en Inglaterra: estar en Londres bastará para estar en Inglaterra. Pero estar en Londres claramente no es una condición necesaria para estar en Inglaterra, ya que hay muchas maneras de estar en Inglaterra sin estar en Londres. Por la misma razón, estar en Inglaterra no es una condición suficiente para estar en Londres.

La teoría del parecido toma la representación pictórica como dependiente del parecido entre la imagen y lo que representa. Expresemos esta dependencia más precisamente en términos de condiciones necesarias y suficientes: una imagen (llámese  $P$ ) representa algo (llámese  $X$ ) si y sólo si  $P$  se parece a  $X$ . O sea que un parecido entre  $P$  y  $X$  es tanto necesario como suficiente para que  $P$  represente  $X$ .

Esta manera de plantear la teoría de la semejanza es ciertamente más precisa que nuestra inicial formulación vaga. Pero, desgraciadamente, expresarla de esta manera más precisa sólo muestra sus problemas. Tomemos primero la idea de que el parecido podría ser una condición suficiente. Y ni siquiera sólo parecido pictórico.

Decir que el parecido es suficiente para la representación es decir esto: si  $X$  se parece a  $Y$ , entonces  $X$  representa  $Y$ . La primera cosa que debe llamarnos la atención es que “se parece” es algo vago. Pues, en un sentido, casi todo se parece a todo lo demás. Éste es el sentido en el cual el algo debe precisamente tener algún rasgo en común con dicha cosa. Así, en este sentido, no sólo se parecen mi padre y mi madre, porque yo me parezco a ellos, sino que también se parecen a mi mesa —mi mesa y yo somos ambos objetos— y al número 3: son ambos objetos de una clase o de otra. Pero no soy una representación de ninguna de estas cosas.

Tal vez necesitemos angostar los modos o aspectos según los cuales algo se parece a algo más si queremos que el parecido sea base de la representación. Pero nótese que no ayuda si decimos que si  $X$  se parece a  $Y$  *en algún aspecto* entonces representa  $Y$ . Pues me parezco a mi padre según ciertos aspectos —digamos rasgos de carácter—, pero esto no me convierte en una representación suya y, evidentemente, no queremos añadir que  $X$  debe parecerse a  $Y$  y a esos aspectos según los cuales  $X$  *representa*  $Y$ , ya que esto haría la teoría del parecido circular y poco informativa: si  $X$  se parece a  $Y$  y a estos segundos aspectos en los cuales  $X$  representa  $Y$ , entonces  $X$  representa  $Y$ . Esto puede ser cierto, pero escasamente será un análisis de la noción de representación.

Hay otro problema en que el parecido sea una condición suficiente. Supóngase que especificamos algunos aspectos según los cuales algo se parece a otra cosa: un retrato de Napoleón, por ejemplo, podría parecerse a Napoleón en la expresión facial, en las proporciones del cuerpo, la posición característica del brazo, y así sucesivamente. Sin embargo, parece ser un hecho evidente acerca del parecido el que si  $X$  se parece a  $Y$  entonces  $Y$  se parece a  $X$ . (Los filó-

sofos plantean esto diciendo que el parecido es una relación *simétrica*.) Si me parezco a mi padre en algunos aspectos, entonces mi padre se parece a mí en ciertos aspectos. Esto no conduce a representación. Si la imagen se parece a Napoleón entonces Napoleón se parece a la imagen. Napoleón no representa la imagen. Así que no puede ser suficiente el parecido para la representación pictórica si hemos de evitar hacer que cualquier objeto sea él mismo una representación pictórica de su imagen.

Finalmente, debemos considerar el hecho evidente de que todo se parece a sí mismo. (Los filósofos plantean esto diciendo que el parecido es *reflexivo*.) Si se supone que el parecido es suficiente condición para la representación, entonces se sigue que todo se representa a sí mismo. Esto es absurdo. No es satisfactoria una teoría de la representación pictórica que convierta *todo* en imagen de sí misma. Esto trivializa por completo la idea de representación pictórica.

Así, la idea de que el parecido sería una condición suficiente de representación pictórica es algo sin esperanza.<sup>3</sup> ¿Significa esto que la teoría del parecido falle? Todavía no, pues la teoría del parecido podría decir que el parecido no es una condición suficiente, sino una condición necesaria. Esto es, que si una imagen *P* representa *X*, entonces *P* se parecerá a *X* en ciertos aspectos, aunque no viceversa. ¿Qué haremos con esta sugerencia?

Tal como se presenta, parece muy plausible. Si un retrato representa a la reina, entonces de seguro debe de parecerse en algún aspecto. Después de todo, puede ser que sea una buena imagen para ser un “buen parecido”. Sin embar-

<sup>3</sup> Véase Nelson Goodman, *Languages of Art* (Indianapolis, Hackett, 1976), cap. 1.

go, hay problemas con esta idea también, pues una imagen puede ciertamente representar a algo sin parecersele mucho. Abundante arte del siglo xx es representacional; pero esto no quiere decir que se base en parecido (considérense los cuadros cubistas). Las caricaturas y los dibujos esquemáticos, al igual que figuras de palos, tienen muy poco parecido en común con las cosas que representan. Sin embargo no acostumbramos preocuparnos reconociendo lo que representan. Una caricatura de la reina puede parecersele mucho menos que un dibujo detallado de alguien diferente. Sin embargo, la caricatura sigue siendo una imagen de la reina.<sup>4</sup>

O sea que ¿cuánta semejanza se requiere para satisfacer la condición necesaria de la representación? Tal vez podría responderse que todo lo que se necesita es que haya *alguna* semejanza, por escasa que sea, entre la imagen y lo que representa. Tal vez el parecido puede tomarse con tanta soltura, que incorpore la representación implicada en los cuadros cubistas. Esto está muy bien, pero ahora la idea de parecido no funciona tanto en la teoría como antes. Si una imagen esquemática (digamos de las usadas en algunas corporaciones en sus logos) necesita parecerse a la cosa que representa sólo de una manera mínima, entonces es difícil ver cuánto es explicado diciendo que “si una imagen representa *X*, debe parecerse a *X*”. Así, incluso cuando una imagen se parece a aquello que representa, debe haber factores diferentes del parecido que entran en la representación y la hacen posible.

<sup>4</sup> Como dice Wittgenstein: “No es la semejanza lo que hace de la imagen un retrato (pudo ser notablemente parecido a una persona, y sin embargo ser un retrato de alguien a quien se parece menos)”. *Philosophical Grammar* (Oxford, Blackwell, 1974), § v.

No estoy negando que las imágenes a menudo se parecen a lo que representan. Evidentemente lo hacen, y esto puede ser parte de lo que las convierte en imágenes (en oposición a las oraciones, las gráficas o los diagramas). Lo único que estoy poniendo en tela de juicio es si la idea de parecido puede *explicar* mucho acerca de cómo representan las imágenes. La idea de que el parecido es una condición necesaria de la representación pictórica bien puede ser cierta, pero la cuestión es “¿qué más hace que una imagen represente como lo hace?”<sup>5</sup>

Un punto que requiere ser subrayado aquí es que las imágenes a menudo necesitan interpretación. Por ejemplo, en el *Juicio Final* de Miguel Ángel, en la Capilla Sixtina, vemos a las almas en el infierno luchando angustiadas al ver su final, con la figura monumental del Cristo encima, alzando la mano para juzgar. ¿Por qué no vemos las almas bienvenidas de las profundidades por el Cristo benévolo, con la mano alzada estimulando amistosamente: “hey, vengan, que aquí está más fresco”? (Recuérdese la imagen de la placa metálica del *Pioneer 10* con su mano alzada.) Pues bien, podríamos, pero no es así. La razón es que vemos la imagen a la luz de ciertos supuestos que hacemos al respecto, a los que podríamos vagamente llamar el “contexto” de la imagen. Sabemos que la imagen es una imagen del Juicio Final, y que en este juicio algunas almas son sentenciadas a la condenación eterna, con el Cristo como juez, y así sucesivamente. Esto es parte de por qué vemos la imagen de la manera como la vemos: la interpretamos.

Podemos situar las cosas mediante un ejemplo de Witt-

<sup>5</sup> Aunque Goodman arguye que no es ni siquiera necesario; véase *Languages of Art*, cap. 1.

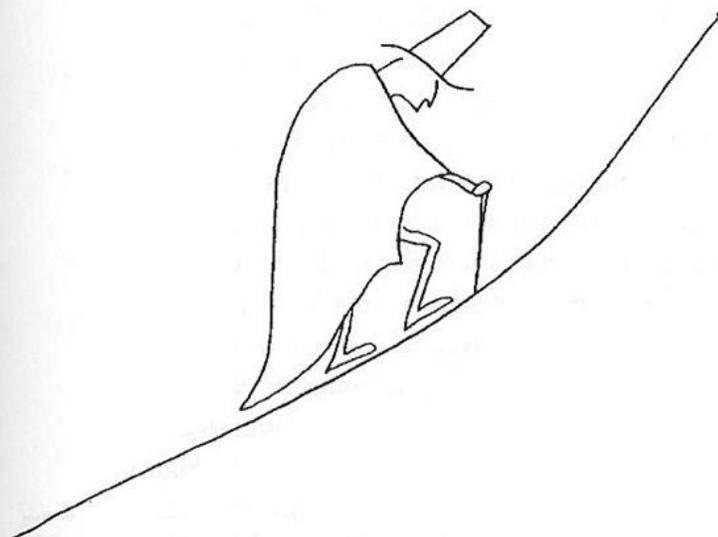


FIGURA 1.1. *Viejo con un bastón*

genstein.<sup>6</sup> Imaginemos la imagen de un hombre con un bastón subiendo una cuesta (véase la figura 1.1). ¿Qué hace que ésta sea la representación de un hombre subiendo por una cuesta, más bien que un hombre deslizándose suavemente cuesta abajo? Nada que haya en la imagen. Es porque estamos acostumbrados a nuestra experiencia cotidiana, y por la clase de contexto en la cual estamos acostumbrados a ver semejantes imágenes, por lo que vemos la imagen de un modo más que de otro. Tenemos que interpretar la imagen a la luz del contexto: la imagen no se interpreta sola.

No voy a perseguir la teoría del parecido o la interpretación de las imágenes en mayor detalle. La menciono aquí para ilustrar cuán poco la idea de parecido nos habla de representación pictórica. Lo que quiero hacer ahora es considerar brevemente la segunda cuestión que planteé al final

<sup>6</sup> *Philosophical Investigations*, p. 54.

de la última sección y aplicarla a la representación pictórica. Podríamos plantear la cuestión así: supongamos que tenemos una teoría completa de la representación pictórica. ¿Sería entonces posible para todas las demás formas de representación ser explicadas en términos de representación pictórica?

La respuesta es “no”, por múltiples razones. Una razón a la cual ya hemos echado una ojeada: las imágenes a menudo requieren ser interpretadas y no servirá el decir que la interpretación debiera ser otra imagen, ya que ésta podría requerir interpretación también. Sin embargo, aunque la respuesta sea negativa, podemos aprender algo acerca de la naturaleza de la representación aprendiendo acerca de las limitaciones de la representación pictórica.

Un ejemplo sencillo puede ilustrar el punto. Supóngase que le digo a usted: “Si no llueve esta tarde, saldremos a pasear”. Esto es una oración bastante sencilla, una representación lingüística. Pero supóngase que queremos explicar *toda* la representación en términos de representación pictórica; requeriríamos poder expresar la representación lingüística en términos de imágenes. ¿Cómo lograríamos esto?

Bien, quizá pudiéramos dibujar una imagen de una escena no lluviosa con usted y yo paseando en ella. ¿Cómo representamos la idea de “esta tarde”? No podemos poner un reloj en la imagen, pues hay que recordar que estamos tratando de reducir toda representación a imágenes, y un reloj no representa el tiempo pintándolo. (La idea de “pintar” el tiempo en realidad tiene poco sentido.)

Y hay la razón adicional de por qué esta primera imagen no podría ser cierta: es sólo una imagen de usted y yo caminando en una zona sin lluvia. Lo que quisimos expresar era una combinación particular y la relación entre dos ideas:

primero, *no* está lloviendo y, segundo, usted y yo salimos a pasear. Así, tal vez debiéramos trazar dos imágenes: una de la escena libre de lluvia y usted y yo caminando. Esto no puede ser tampoco correcto, pues ¿cómo podría este par de imágenes expresar la idea de que *si* no llueve, *entonces* saldremos a pasear?, ¿por qué no debieran tomarse las imágenes como representación sencillamente de una escena sin lluvia y usted y yo saliendo a pasear? O ¿por qué no representa la idea de que *o bien* saldremos de paseo *o* no lloverá? Cuando tratamos de representar la diferencia entre ... *y* ..., *si...* *entonces...*, *y...* *o bien...* en imágenes, dibujamos un blanco completo. Parece simplemente no haber otra manera de hacerlo.

Una cuestión importante que las imágenes no pueden hacer, pues, es representar ciertas clases de relaciones entre ideas. No pueden representar, por ejemplo, las relaciones que expresamos utilizando las palabras *si...*, *entonces...*, ... *y...*, *o bien...* *o...* *y no*. (¿Por qué *no*? Pues bien, la imagen de la escena sin lluvia puede igualmente ser una imagen de una escena soleada, ¿cómo podemos expresar pictóricamente la idea de que la escena es una escena *sin* lluvia? Tal vez dibujando lluvia y poniendo una cruz —como en un signo de “No fumar”— pero otra vez estamos usando algo que no es una imagen: la cruz.) Por esta razón cuando menos, es imposible explicar o reducir otras formas de representación a la representación pictórica.

#### REPRESENTACIÓN LINGÜÍSTICA

Una imagen a veces puede valer por mil palabras, pero mil imágenes no pueden representar algunas de las cosas que

podemos representar usando palabras y oraciones. Así, ¿cómo podemos representar cosas usando palabras y oraciones?

Una idea natural es ésta: “Las palabras no representan cosas de ninguna manera natural; más bien representan por *convención*. Hay una convención entre los hablantes de un lenguaje de que las palabras que usan significarán la misma cosa para el uno y el otro; cuando los hablantes convengan o converjan en sus convenciones, lograrán comunicarse; cuando no, no lo harán”.<sup>7</sup>

Es difícil negar que lo que las palabras representan es al menos en parte una cuestión de convención. ¿Cuál es la convención, exactamente? Considérese la palabra “perro”. ¿Es la idea de que hay una convención entre los hablantes de nuestro idioma para usar la palabra “perro” para representar perros, y sólo perros (mientras se trate de hablar literalmente y hablar con verdad)? Entonces es difícil ver cómo la convención puede *explicar* la representación, según afirmamos la convención como “una convención de usar la palabra ‘perro’ para *representar* perros”. Como la convención se enuncia usando la idea de representación, la da por descontado: no puede explicarla. (De nuevo, mi punto en cuestión no es que la convención no esté implicada en la representación lingüística; la cuestión es si el recurso a la convención puede explicarla de por sí.)

Un pensamiento igualmente natural es que las palabras representan por estar enlazadas convencionalmente a las

<sup>7</sup> Esto evidentemente es una manera muy sencilla de plantear la cuestión. Para más acerca de la convención, véase David Lewis, *Convention* (Oxford, Blackwell, 1969). Sobre el escepticismo acerca del papel de la convención en el lenguaje, véase Donald Davidson, “Communication and Convention”, *Inquiries into Truth and Interpretation* (Oxford, Oxford University Press, 1984).

*ideas* que los pensadores intentan expresar usando estas palabras. La palabra “perro” expresa la idea de un perro por medio de una convención que enlaza la palabra a la idea. Esta teoría tiene una historia filosófica especial: algo parecido a ella se remonta al menos a Thomas Hobbes (1588-1679), y especialmente John Locke (1632-1704), que resumieron el punto de vista diciendo que las palabras son “marcas sensibles de ideas”.<sup>8</sup>

¿Qué son ideas? Algunos filósofos han sostenido que son algo como las imágenes mentales: imágenes en la mente. Así, cuando uso la palabra “perro”, esto se correlaciona con una imagen mental, en mi mente, de un perro. Una convención asocia la palabra “perro” con la idea que está en mi mente, y es en virtud de esta asociación como la palabra representa perros.

Esta teoría plantea muchos problemas. Sin ir más lejos, ¿es la imagen de mi mente una imagen de un perro particular, digamos Firuláis? En tal caso, ¿suponemos que la palabra “perro” significa *perro*, más bien que *Firuláis*? Además, es difícil suponer qué imagen de “perrería” en general sería.<sup>9</sup> Y aun si la teoría de las imágenes mentales de ideas de alguna manera puede dar razón de este problema, encontrará el problema mencionado al final de la última sección. Aunque muchas palabras pueden asociarse con imágenes mentales, otras muchas no: éste era el problema que teníamos al tratar de explicar *y, o, no y sí* en términos de imágenes.

Sin embargo, tal vez no todas las ideas son imágenes mentales —a menudo pensamos en palabras, por ejemplo,

<sup>8</sup> John Locke, *An Essay Concerning Human Understanding* (1689), libro III, cap. 2, § 1.

<sup>9</sup> Véase la crítica por George Berkeley de la doctrina de Locke de las ideas abstractas en *Principles of Human Knowledge* (1710).

y no en absoluto en imágenes—. En tal caso, las críticas a los dos últimos párrafos no dan en el blanco. De modo que pongamos a un lado la teoría de que las ideas son imágenes mentales, y nada más consideremos la pretensión de que las palabras representan expresando ideas, cualesquiera ideas que sean.

Esta teoría no recurre a una “convención para *representar* perros”, de modo que no es vulnerable a la misma crítica que la teoría anterior. Pero no puede, por supuesto, explicar la representación, porque atrae ideas, y ¿qué son las ideas sino otra forma de representación? Una idea de perro representa perros tanto como la palabra “perro”; de modo que estamos en efecto recurriendo a una clase de representación (la idea) para explicar otra clase (la palabra). Esto está muy bien, pero si queremos explicar la representación en general, entonces también necesitamos explicar cómo representan las *ideas*.

Tal vez se piense que esto es preguntar demasiado. Tal vez no necesitamos explicar cómo las ideas representan. Si explicamos cómo las palabras representan, asociándolas con ideas, y explicamos también cómo las imágenes pueden ser interpretadas en términos de las ideas que la gente asocia con ellas en sus mentes, quizás podamos detenernos ahí. En resumidas cuentas, no podemos explicarlo todo y tenemos que dar algo por sabido. Así que ¿por qué no tomar las capacidades representacionales de las ideas como cosa establecida?

Pienso que esto es insatisfactorio. Si estamos contentos con dar por sabidos los poderes representacionales de la mente, entonces ¿por qué no dar un paso atrás y tener por sabidos los poderes del lenguaje? Porque no es tan bien entendida la mente como el lenguaje, de hecho en filosofía lo

contrario es lo probablemente cierto. Las ideas, los pensamientos y los fenómenos mentales generalmente parecen hasta más misteriosos que las palabras y las imágenes. De modo que, si acaso, esto sugeriría que debiéramos explicar las ideas en términos del lenguaje, antes que a la inversa. Sin embargo, no pienso que podamos hacer esto. Así es que necesitamos explicar la naturaleza representacional de las ideas.

Antes de pasar a discutir las ideas y la representación mental, quisiera dejar en claro lo que estoy diciendo acerca de la representación lingüística. No estoy diciendo que las nociones que mencioné —de convención, o de palabras que expresan ideas— sean las únicas opciones para una teoría del lenguaje. No, en absoluto. Las introduje sólo como ejemplos de cómo una teoría de la representación lingüística deberá, a fin de cuentas, echar mano de una teoría de la representación mental. Algunas teorías del lenguaje negarán esto, pero ignoraremos estas teorías aquí.<sup>10</sup>

La conclusión de esta discusión es que las palabras, igual que las imágenes, no representan en sí mismas (“intrínsecamente”). Necesitan interpretación, necesitan una interpretación que se les asigne de alguna manera. ¿Cómo podemos explicar esto? La respuesta natural, pienso, es que la interpretación es algo que la *mente* concede a las palabras. Palabras e imágenes ganan las interpretaciones que tienen, y por lo tanto representan lo que representan, en virtud de los estados mentales de quienes las usan. Estos estados menta-

<sup>10</sup> Véase, por ejemplo, el intento de Davidson de dilucidar el significado lingüístico en términos de verdad: *Inquiries into Truth and Interpretation* (Oxford, Oxford University Press, 1984). Vista de conjunto, Barry C. Smith, “Understanding Language”, *Proceedings of the Aristotelian Society*, 92 (1992).

les son representacionales también. De manera que para comprender cabalmente la representación lingüística y pictórica tenemos que comprender la representación mental.

#### REPRESENTACIÓN MENTAL

Así que ¿cómo representa la mente algo? Hagamos la cuestión un poco más fácil de manejar preguntando cómo los *estados* individuales de la mente representan algo. Por un “estado de la mente”, o “estado mental”, quiero decir aquí algo como una creencia, un deseo, una esperanza, un miedo, una conjetura, una expectación, una intención, una percepción, y así por el estilo. Pienso que todos estos son estados de la mente que representan el mundo de alguna manera. Esto requerirá un poco de explicación.

Cuando digo que esperanzas, creencias, deseos y demás representan el mundo, quiero decir que toda esperanza, creencia o deseo *se dirige a* algo. Si se espera, hay que esperar *algo*; si se cree se debe creer en *algo*; si se desea debe desearse *algo*. No tiene sentido suponer que una persona podría simplemente esperar, sin esperar algo; creer sin creer algo; o desear sin desear algo. Lo que usted crea o desee es lo que es representado por su creencia o deseo.

Necesitaremos un término general conveniente para estados de la mente que representen el mundo, o un aspecto del mundo. Usaré el término “pensamiento”, ya que parece ser el término más general y neutro perteneciente al vocabulario mental cotidiano. De aquí en adelante en este libro usaré el término “pensamiento” para aludir a todos los estados mentales representacionales. Así, los estados de creencia, deseo, esperanza, amor y demás son todos pensamientos en

mi sentido, ya que todos representan cosas. (Cuándo todos los estados mentales son pensamientos en este sentido, es cuestión que dejaré hasta el final del capítulo.)

¿Qué podemos decir en general acerca de cómo los pensamientos representan? Comenzaré con pensamientos que son de particular interés; esos pensamientos que representan *situaciones* (o tratan de ellas). Cuando espero que haya sopa de mariscos en el menú en mi restaurante esta noche, estoy pensando en diversas cosas: la sopa, el menú, mi restaurante favorito, esta noche. Pero no estoy pensando sólo en estas cosas de un modo casual o desconectado: estoy pensando en cierto hecho o *situación* posibles: la situación en la cual la sopa está en el menú de mi restaurante favorito esta noche. Es una variante inocua de esto el decir que mi estado de esperanza *representa* esta situación.

Sin embargo, considérese un pensamiento diferente que podría tener: la *creencia* de que hay sopa de mariscos en el menú esta noche. Este estado mental no representa la situación del mismo modo como la esperanza lo hace. Cuando creo que hay sopa de mariscos en el menú esta noche (porque he pasado por el restaurante y leído el menú), tomo la situación en cuestión como un hecho: tomo como un hecho acerca del mundo que hay sopa de mariscos en el menú esta noche. Pero cuando espero, no lo esperaré como un hecho acerca del mundo; más bien me gustaría que fuera un hecho que hubiese sopa de mariscos en el menú esta noche.

Así que hay dos aspectos de estos pensamientos: hay la “situación” representada y hay lo que podríamos llamar (a falta de una palabra mejor) la *actitud* que asumimos ante la situación. La idea de diferentes actitudes ante situaciones es ilustrada mejor mediante ejemplos.

Considérese la situación de que yo visite Budapest. Pue-

do esperar que visitaré Budapest, y puedo creer que he visitado Budapest. Todos estos pensamientos tratan o representan la misma situación —mi visita a Budapest— pero las actitudes asumidas ante esta situación son muy diferentes. La cuestión por lo tanto surge acerca de lo que hace diferentes a estas distintas actitudes; pero por el momento sólo me importa distinguir la situación representada a partir de la actitud asumida ante ella.

Así como la misma situación puede ser tema de diferentes actitudes, así el mismo género de actitud puede concernir a muchas situaciones diferentes. Realmente creo que visitaré Budapest pronto, y también creo que mi restaurante favorito no tiene sopa de mariscos en el menú de esta noche, y creo otras innumerables cosas. Creencias, esperanzas y pensamientos de este género pueden ser captados únicamente especificando: *a*) la actitud en cuestión (creencia, esperanza, expectación, etc.) y *b*) la situación representada. (Debe también notarse, de paso, que muchas actitudes llegan en grados: uno puede querer algo más o menos intensamente; y creer algo con más o menos convicción; pero esta complicación no afecta al cuadro general.) En general podemos describir estas clases de pensamientos esquemáticamente como sigue. Si “*A*” es la persona que se halla en el estado mental, “ $\psi$ ” representa la actitud (la letra psi, por “psicológico”) y “*S*” ocupa el lugar de la situación representada, la mejor descripción será de la siguiente forma:

Un  $\psi$  que *S*

Por ejemplo, Vladimir (*A*) cree ( $\psi$ ) que está lloviendo (*S*); Renata (*A*) espera ( $\psi$ ) que visitará Rumania (*S*), y así sucesivamente.

Bertrand Russell (1872-1970) llamó a los pensamientos que pueden ser pescados de esta manera “actitudes proposicionales”, y esta denominación ha corrido con suerte.<sup>11</sup> Aunque puede resultar más bien oscuro al primer golpe de vista, el término “actitud proposicional” describe la estructura de estos estados mentales sumamente bien. Ya he explicado el término “actitud”. Lo que Russell quiere decir con “proposición” es algo parecido a lo que estoy llamando “situación”: es aquello hacia lo cual se dirige la atención (así, una proposición en este sentido no es un fragmento del lenguaje). Una actitud proposicional es por lo tanto cualquier estado mental que pueda describirse en el estilo “*A*  $\psi$  que *S*”.

Otro fragmento de terminología que ha sido casi universalmente adoptado es el término “contenido”, usado donde Russell usó “proposición”. De acuerdo con esta terminología, cuando creo que hay cerveza en el refrigerador, el *contenido* de mi creencia es que *hay cerveza en el refrigerador*. Y análogamente con deseos, esperanzas y así sucesivamente, éstas son actitudes diferentes, pero todas tienen “contenido”. Qué es exactamente “contenido” y qué es para un estado mental tener “contenido” (o “contenido representacional”) son cuestiones que retornarán durante el resto de este libro, especialmente en el capítulo v. En la filosofía del momento, el problema de la representación mental se expresa a menudo como: “¿Qué es para un estado mental tener contenido?” Por el momento, podemos pensar en el contenido de un estado mental como aquello que distingue estados que implican la misma actitud el uno para el otro.

<sup>11</sup> Russell usó la expresión en *The Analysis of Mind* (Londres, George Allen and Unwin, 1921), cap. 12. Para una colección de lecturas, véase Nathan Salmon y Scott Soames (eds.), *Propositions and Attitudes* (Oxford, Oxford University Press, 1988).

Diferentes creencias se distinguen una de otra (o, en terminología filosófica, son “individuadas”) por sus diferentes contenidos. Así son los deseos; y así con todas las actitudes.

Me he concentrado en la idea de una actitud proposicional, porque los pensamientos de esta forma se volverán muy importantes en el capítulo siguiente. Y aunque todas las actitudes proposicionales sean pensamientos (por definición), es importante recalcar que no todos los pensamientos (en mi sentido) son actitudes proposicionales, esto es, no todos los estados mentales representacionales pueden caracterizarse en términos de actitudes ante situaciones. Tómese el amor, por ejemplo. El amor es un estado mental representacional: no se puede amar sin amar algo o a alguien. Sin embargo, el amor no es (siempre) una actitud ante una situación; el amor puede ser una actitud hacia una persona, un lugar o una cosa. El amor no puede ser descrito en el estilo de “ $A \psi$  que  $S$ ” (inténtese y véase). En mi terminología, entonces, el amor es una clase de pensamiento, pero no una actitud proposicional.<sup>12</sup>

Otro interesante ejemplo es el deseo. ¿Es éste una actitud hacia una situación? Vistas las cosas, no lo es. Supóngase que deseo una taza de café; mi deseo es por una cosa, una taza de café, no por ninguna situación. En la superficie, pues, el deseo se asemeja al amor. Muchos filósofos piensan que esto es engañoso, y que describe poco de un deseo para tratarlo como una actitud hacia una cosa. La razón es que una descripción más exacta del deseo es que es un deseo de que se dé una situación determinada: la situación en la cual *yo tengo una taza de café*. Todos los deseos, según se sostiene,

<sup>12</sup> Para más acerca de este tema, véase Tim Crane, *Elements of Mind* (Oxford, Oxford University Press, 2001), § 34.

ne, son realmente deseos de *que esto y lo otro*, donde “esto y lo otro” es una especificación de una situación. El deseo, a diferencia del amor, es una actitud proposicional.

Ahora, denominando “pensamientos” a los estados mentales representacionales, no quiero dar a entender que estos estados sean necesariamente conscientes. Supóngase que Edipo realmente no desea matar a su padre ni casarse con su madre. Entonces, según el criterio esbozado antes ( $A \psi$  que  $S$ ), estos deseos cuentan como actitudes proposicionales y por lo tanto pensamientos. No obstante, no son pensamientos conscientes.

Podría parecer extraño distinguir entre pensamiento y conciencia de esta manera. Para justificar la distinción, necesitamos una breve digresión preliminar sobre el tema turbio de la conciencia; un tratamiento cabal de este asunto habrá de esperar hasta el capítulo vi.

#### PENSAMIENTO Y CONCIENCIA

La conciencia es lo que hace que nuestras vidas despiertas parezcan ser como son, y puede sostenerse que también es la fuente última de todo valor del mundo: “Sin esta iluminación interna —dijo Einstein al filósofo Hebert Feigl— “el universo no sería sino un montón de tierra”.<sup>13</sup> Sin embargo, a pesar de la importancia de la conciencia, quiero distinguir ciertas cuestiones acerca del pensamiento a partir de preguntas sobre la conciencia. En cierta medida, estas cuestiones son independientes una de otra.

<sup>13</sup> Citado en H. Feigl, *The “Mental” and the “Physical”* (Minneapolis, University of Minnesota, 1967), p. 138.