

# Interpretando la Paradoja de Moore: la irracionalidad de una oración mooreana<sup>1</sup>

(*Interpreting Moore's Paradox: the irrationality of a Moorean sentence*)

Cristina BORGONI

Recibido: 17.03.2008

Versión final: 26.05.2008

BIBLID [0495-4548 (2008) 23: 62; pp. 145-161]

**RESUMEN:** Este trabajo ofrece una lectura de la Paradoja de Moore que pone énfasis en su relevancia para nuestra comprensión de la racionalidad y de la interpretación lingüística. Mantiene que las oraciones que dan origen a la paradoja no necesitan entenderse en términos de ausencia de una contradicción, sino más bien en términos de ausencia de racionalidad, entendida esta como un término más amplio que el de coherencia y consistencia lógica. Se defenderá tal posición por medio de tres tesis, dos de las cuales se derivan de los enfoques dominantes (aunque insuficientes) a la paradoja: el de Moore, el de Wittgenstein y el de Shoemaker.

**Descriptores:** Paradoja de Moore, racionalidad, interpretación radical, unidad personal, externalismo y la perspectiva de primera persona.

**ABSTRACT:** *This paper offers an interpretation of Moore's Paradox that emphasizes its relevance for our understanding of rationality and linguistic interpretation. The sentences that originate the paradox do not need to be thought of in terms of the absence of a contradiction, but in terms of absence of rationality, where rationality is understood as a broader notion than coherence and logical consistency. This is defended through three theses, two of which stem from the dominant (but insufficient) approaches to the paradox: Moore's, Wittgenstein's and Shoemaker's.*

**Keywords:** *Moore's Paradox, rationality, radical interpretation, personal unity, externalism and first person perspective.*

## Introducción

Imaginemos que llueve y que estamos una amiga y yo frente a una ventana donde se ve claramente tal situación. Ella dice: “Está lloviendo”. Yo la escucho y entiendo que quiere decir algo respecto al tiempo; quiere, por ejemplo, llamarme la atención sobre un evento que estamos presenciando y sobre el que, de alguna manera, podemos pensar. En seguida afirma, “pero no creo que esté lloviendo”. ¿Qué puedo decir en respuesta? Por un lado, soy capaz de ver que llueve y sé que se equivoca cuando afirma que no cree en ello. Pero, por otra parte, al haber dicho “llueve” al principio, me deja sin espacio para darle cualquier razón que la convenza de lo contrario con relación a la segunda parte de su enunciado. Parece que solamente me resta insistirle “¿Qué?” A pesar de tal impase, tengo que aceptar que mi amiga ha proferido una oración gramaticalmente correcta.

---

<sup>1</sup> Versiones anteriores de este trabajo fueron presentadas y discutidas en la Universidad de Granada, en marzo de 2007, en el marco del V Seminario de Trabajos en Construcción, y en Barcelona, en septiembre de 2007, en el marco del V congreso de la SEFA. Agradezco el interés y los comentarios allí recibidos. Agradezco especialmente a Manuel de Pinedo García, a Nefthalí Villanueva Fernández, a María del Mar Muriana López, a Miguel Ángel Pérez y a los dos evaluadores anónimos.

La discusión sobre la Paradoja de Moore se establece en contextos similares a este, donde existe un enunciado que, en principio no tiene ningún problema formal, pero resulta ser un enunciado absurdo. Tradicionalmente, la paradoja se caracteriza por un consenso sobre el carácter absurdo de tales oraciones, a la vez que no existe ninguna contradicción en la oración en sí misma. Por esto, la discusión se convierte en una búsqueda de lo que podría ser una contradicción disfrazada.

En este texto buscaré ofrecer una lectura de la paradoja que ponga énfasis en su relevancia para nuestra comprensión de la racionalidad y de la interpretación lingüística. Defenderé que tales oraciones paradójicas no necesitan entenderse en términos de ausencia de una contradicción sino más bien, en términos de ausencia de racionalidad. Lo haré por medio de la defensa de tres tesis, dos de las cuales se derivan de los enfoques dominantes (para mí, insuficientes) a la paradoja: el de Moore, el de Wittgenstein y el de Shoemaker<sup>2</sup>. Con esto, intentaré mostrar que tal paradoja no necesita verse como un problema encerrado en sí mismo, sino que el camino para solucionarla pasa por cuestiones importantes concernientes a cómo funciona nuestra interpretación lingüística, a la auto-atribución de conocimiento y al carácter externo de lo mental.

Mi estrategia será: primero, estimular el carácter paradójico de las oraciones en juego, luego presentar las tres explicaciones dominantes acerca de la paradoja y por último, exponer mi propuesta y los temas que están involucrados en ella.

### 1. *Qué hay de paradójico en las “oraciones mooreanas”*

Consideremos las siguientes oraciones:

- i.  $p$
- ii. Creo que  $p$
- iii. No creo que  $p$
- iv. Creo que no  $p$

como expresión de situaciones en el mundo. Puede pasar que lo que (i) expresa ocurra conjuntamente con cada uno de los casos expresados por las tres siguientes oraciones:  $p$  podría ser el caso y además de esto, yo podría creer que  $p$ ; otra situación sería aquella en la que  $p$  fuera el caso y yo ignorara tal cosa; y finalmente,  $p$  podría ser el caso y a la vez yo podría equivocadamente creer en la negación de  $p$ . A pesar de la plausibilidad de tales situaciones, enunciar (i) más (iii) o (i) más (iv) como:

1.  $p$ , pero no creo que  $p$  (i & iii)

y

2.  $p$ , pero creo que no  $p$  (i & iv)

produciría, en ambos casos, enunciados muy extraños.

---

<sup>2</sup> Agradezco a un evaluador anónimo el hacerme ver la importancia de la solución de Hintikka (1962). Sin embargo, he optado por hacer referencias a ella en distintas partes de este trabajo en lugar de tratarla como una cuarta lectura de la paradoja. Como señalaré, las ideas de Hintikka tienen aspectos en común con algunas de las soluciones que serán estudiadas, incluyendo la que yo propondré.

Es aquí donde la “paradoja de Moore” aparece y por esto, a partir de ahora, las oraciones de los tipos (1) y (2) pasarán a llamarse “oraciones mooreanas” (OM). El mundo puede fácilmente aceptar que algo sea el caso sin que yo crea en ello, pero tenemos el deber de recusar las palabras de alguien que enuncia un hecho y a la vez enuncia su creencia en lo contrario. Del mismo modo que debemos rechazar que alguien afirme algo a la vez que afirma la ausencia de su creencia en tal cosa. Dicho de una manera más tradicional, no hay ninguna contradicción entre (i) y (iii) ni tampoco entre (i) y (iv) pero, aún así, cuando tales conjunciones están afirmadas en primera persona se vuelven inaceptables, porque son absurdas.

Según Williams (2006, p. 227) Moore habría tenido el cuidado de distinguir un absurdo de una paradoja. En el caso de las oraciones mooreanas, lo absurdo sería que yo afirmara tales oraciones. Lo paradójico sería que el absurdo coexistiera con la ausencia de cualquier contradicción en mis palabras.

A lo largo del texto voy a mantener que tales oraciones son, de hecho, inaceptables, porque dejan de responder a un comportamiento racional y no por ser un equívoco lógico. Pero, por ahora es suficiente con que sigamos delimitando el tipo de oraciones con las que estamos tratando. Es interesante señalar la peculiaridad de (1) y (2) frente a otras oraciones que tienen alguna similitud con una OM, pero no lo son, como las siguientes:

3.  $p$ , pero ella no cree que  $p$
4.  $p$ , pero yo creía que no  $p$
5.  $p$ , pero sé que no  $p$

No hay nada de contradictorio ni tampoco de paradójico en enunciar (3) o (4). Ambas podrían ser incluso verdaderas.

El aire paradójico de “ $p$ , pero no creo que  $p$ ” en contraste con “ $p$ , pero ella no cree que  $p$ ” (o incluso “ $p$ , pero yo creía que no  $p$ ”) nos invita a reconocer la diferencia entre puntos de vista de primera y de tercera persona. Aunque (1) y (3) tengan la misma forma lógica ( $p \ \& \ \sim B_s p$ ), la diferencia entre que una sea paradójica y la otra no, se halla en quién es “ $S$ ”; quién es el sujeto que cree en  $p$ <sup>3</sup>. En (1) hablo de mí misma mientras que en (3) hablo de otra persona. Y, con frecuencia, nos apercebimos de la ocurrencia de determinado evento al tiempo que reconocemos que hay personas que no creen que suceda.

Con respecto a (4), la oración es igualmente familiar. Son muchas las veces en que nos damos cuenta de que algo es el caso aunque creyéramos erróneamente en lo contrario. En este caso, la oración la enuncio yo sobre mí misma, pero también desde una perspectiva de tercera persona. Tengo que verme a mí misma en el pasado y comprenderme como poseedora de una creencia que ya no sostengo. Y para esto parece

---

<sup>3</sup> Hintikka (1962, p. 104) llama la atención sobre tal diferencia y reconoce que el pronombre de primera persona “yo” juega aquí un papel especial. Sin embargo, sugiere que no solamente sería preferible explicar la paradoja sin añadir criterios que hicieran referencia al mismo, sino que una explicación de este tipo, “en términos de las características propias de los pronombres de primera persona” estaría en principio mal encaminada.

necesario un cierto alejamiento que se consigue asumiendo una perspectiva de tercera persona sobre mí misma. Si digo (4), al localizar mi creencia en un tiempo distinto, me libro de estar en una situación absurda<sup>4</sup>.

La quinta oración es bastante distinta de las anteriores, porque en ella sí podemos encontrar una contradicción clara entre  $p$  & no  $p$ . Y esto se debe a la facticidad del conocimiento. En el caso (5) podríamos decir que yo, o quien lo afirmara, estaría cometiendo un equívoco porque estaría afirmando una contradicción, pero no tendríamos problemas en explicar el porqué. Así que, no estaríamos ante una oración paradójica. Si alguien conoce  $p$ , implica que  $p$  es el caso. Aquí no hay ningún misterio<sup>5</sup>.

Sería posible establecer otras relaciones entre las oraciones mooreanas y otras que se parecen a ellas pero no lo son. Sin embargo, creo que estas son suficientes para marcar algunas especificidades de una OM, y más que esto, motivar el verlas como oraciones problemáticas. El hecho de que las oraciones mooreanas sean una conjunción entre un enunciado del tipo “ $p$ ” y otro del tipo “ $S$  cree que  $p$ ” (o bien  $S$  no tiene la creencia en  $p$ , o bien  $S$  cree en lo contrario de  $p$ ), y el hecho de que se afirmen por la primera persona y en tiempo presente, son suficientes para que surja la siguiente cuestión: ¿qué tiene de absurdo una OM? En la próxima sección veremos una variedad de respuestas a tal pregunta.

## 2. *Disolviendo la paradoja*

Dar una solución a la presente paradoja es descubrir dónde está lo absurdo de una OM. Para esto, se ha sugerido que debemos localizar el fenómeno que se parece a una contradicción en otra parte que no sea la oración en su forma primitiva. Hay, en realidad, muchas propuestas de solución en la literatura, pero, por lo menos tres me parecen especialmente importantes y serán estas las que consideraré: la primera está inspirada por el mismo Moore, la segunda desarrollada por Wittgenstein, y una tercera por Shoemaker. Las analizaré con la intención de buscar las principales intuiciones que subyacen a cada solución.

### *G. E. Moore*

La paradoja de Moore, incluso la invención del propio término “Moore’s Paradox”, debe su significado filosófico, tal y como lo conocemos y lo discutimos hoy, a Wittgenstein (1953). Según observa Baldwin (1996, p. 226) G. E. Moore habría presentado su paradoja en dos trabajos tardíos y en ninguno de ellos la paradoja era el

<sup>4</sup> Alguien podría contra-argumentar que asumir la perspectiva de tercera persona sobre mí misma no me libra de caer en algún momento en la paradoja de Moore. Tal perspectiva podría tomarse también para hablar sobre mis estados de creencia actuales con la ayuda, por ejemplo, de un psicólogo, y podría así afirmar algo como la oración (1). R. Moran (2001) llama la atención justamente sobre casos como estos en los que podríamos tener una perspectiva de primera persona conflictiva con una de tercera persona sobre nosotros mismos. Estoy totalmente de acuerdo con que casos así podrían darse. Sin embargo, la peculiaridad de (4) está en que el modo como accedo a mí parece estar presente en la propia oración. No hay ningún problema en que yo actualmente crea en  $p$ , aunque no haya creído en el pasado.

<sup>5</sup> T. Williamson (2000) defiende que oraciones como la (5) pueden entenderse como casos de la paradoja de Moore.

tema principal. La utilización de los enunciados que dieron lugar a ella, servía solamente para destacar la distinción entre lo que una persona afirma y lo que implica al hacerlo.

Es posible, sin embargo, encontrar una solución a este problema inspirada en los textos del propio Moore. En su famoso ejemplo: “Fui a la exposición el jueves pasado, pero no creo que fui”<sup>6</sup> (Moore 1942, p. 543), Moore sostiene que mi declaración de que fui a la exposición implica que creo que fui. Y es por esto que sostiene que proférer tal oración sería absurdo.

Una posible interpretación de la idea de que “mi declaración de que fui a la exposición implica que creo que fui” sería mantener que (i) implica lógicamente (ii)<sup>7</sup>. Bajo tal interpretación, si alguien asevera (i), podríamos inferir (ii), permitiéndonos ver la tensión buscada. En el caso de la preferencia de (1)  $\neg p$ , pero no creo que  $p$ — si seguimos tal sentido lógico de implicar, podríamos inferir: (1\*) “creo que  $p$ , pero no creo que  $p$ ”, mientras en el caso de (2)  $\neg p$ , pero creo que no  $p$ — tendríamos: (2\*) “creo que  $p$ , pero creo que no  $p$ ”.

La diferencia entre las oraciones (1) y (2), del mismo modo que en las resultantes (1\*) y (2\*), puede, en principio, considerarse como una diferencia demasiado sutil una vez que utilizamos normalmente la oración (iii) “no creo que  $p$ ” con el sentido de (iv) “creo que no  $p$ ”. Sin embargo, Williams (1979) sugiere que tal diferencia debe tenerse en consideración porque, en el caso de la paradoja de Moore, se establecen dos problemas donde antes solo se veía uno, de manera tal que su solución debe entenderse desde este doble aspecto.

Según Williams (1979, p. 142) el absurdo de (1) se distingue del absurdo de (2) de la siguiente manera: en el primer caso sería absurdo que alguien dijera (1) porque lo que se expresa y se afirma conjuntamente, o sea, (1\*): una creencia de que  $p$  y la ausencia de la creencia de que  $p$ , es *imposible lógicamente*. Mientras que en el caso de (2) lo que se expresa y se afirma conjuntamente, o sea, (2\*): una creencia que  $p$  y una creencia de que no es el caso que  $p$ , es *inconsistente*.

Seguiré considerando (1) y (2) como casos distintos, pero no me comprometeré con las denominaciones señaladas por Williams para cada caso porque estas solamente tienen sentido si aceptamos esta primera interpretación. Una interpretación que, sin embargo, no parece ser la interpretación más fiel a la posición de propio Moore. Aunque Moore no desarrolle una solución sistemática a la paradoja, se puede entender desde sus escritos que el sentido de la implicación a la cual se refiere no es el de implicación lógica, sino más bien una noción informal y cotidiana<sup>8</sup>. En uno de sus trabajos,

<sup>6</sup> Todas las traducciones del inglés son mías a no ser que se indique lo contrario.

<sup>7</sup> Se mantendrá más adelante que tal interpretación no es la interpretación más fiel a la propia posición de Moore. Sin embargo, como se va a discutir también más adelante, la interpretación que sí sería la más fiel no logra una solución satisfactoria en el sentido de encontrar una contradicción en las oraciones en juego. Se considerará inicialmente esta primera interpretación porque además de constituir una lectura posible, parece estar presupuesta por otros filósofos como es el caso de Williams (1979).

<sup>8</sup> Hay razones de peso para entender que cuando Moore habla de implicación no se refiere a la implicación lógica, sino más bien a la implicación epistémica o pragmática (esto es algo en lo que insiste, por ejemplo, Hintikka (1962) en su lectura de la paradoja). Max Black (1952), aunque no utilice términos

Moore (1944, p. 204, citado por Black 1952, p. 26) sostiene que si decimos, por ejemplo, “él no salió” nosotros implicamos que no creemos que él haya salido, aunque esto no haya sido ni afirmado ni pueda seguirse de nada de lo que afirmamos<sup>9</sup>. Según expone, implicamos esto porque llevamos a nuestros oyentes a asumir tal cosa, porque “las personas, en general, no hacen una aserción positiva a no ser que no crean que su opuesto sea verdadero”. Por esta razón, la primera interpretación que prometía una solución rápida a la paradoja no se sostiene. En parte porque no es la posición de Moore, pero principalmente porque no parece ser sencillo mantener que haya una implicación lógica entre (i) y (ii). Si uno puede suponer (ii) cuando alguien asevera (i), esto no es lo mismo que sostener que en todo caso, (i) implica (ii). Y Moore no es ingenuo con respecto a esto.

Nos quedamos entonces con la otra interpretación de la posición de Moore, la que toma el sentido de la implicación entre mi afirmación sobre un evento y mi creencia en tal evento como más débil que el sentido de “implicación lógica”. Linville & Ring (1991, pp. 295-296), teniendo esto en cuenta, señalan sin embargo que solamente en este segundo sentido (el sentido lógico) tendríamos derecho a indicar la existencia de una contradicción en una OM y así solucionar la paradoja. Una vez que la paradoja es puesta en términos de hallar una contradicción, esta segunda interpretación de Moore se muestra insatisfactoria<sup>10</sup>.

tan precisos como los de la pragmática contemporánea, ofrece una lectura sobre el uso que hace Moore de la noción de inferencia que puede resultar esclarecedor: “De estas observaciones sobre Moore, podemos derivar la siguiente explicación del modo según el cual él [Moore] está usando aquí la palabra ‘implica’: Suponga (i) el hablante está usando una expresión,  $E$ , (ii) las personas no usan en general  $E$  sin que alguna proposición relacionada  $pE$  sea verdadera, (iii) las personas, al oír el uso de la expresión  $E$  por el hablante, en general van a creer que  $pE$  es verdadera, y por fin (iv) el hablante sabe todo esto —entonces si las cuatro condiciones se dan, las palabras del hablante pueden ser tomadas como implicando  $pE$ ” (Black, 1952, p. 26).

<sup>9</sup> Aquí se podría recurrir a nociones pragmáticas como es el caso de la noción de “afirmación” entendida como un acto de habla sujeto a determinadas normas, como por ejemplo, la suposición de que quien afirma algo cree lo que afirma. Pero, de nuevo, es importante señalar que leer a Moore en términos de las categorías de la pragmática contemporánea es una entre varias interpretaciones posibles, ya que él no disponía de tales recursos en sus trabajos.

<sup>10</sup> Otra posible salida sería tomar una dirección como la que toma Hintikka (1962), que apela a principios de creer y de saber para explicar el carácter paradójico de una OM. De una cierta manera, Hintikka reconoce la importancia del punto señalado por Moore, de que hay un cierto presupuesto en la comunicación de que las personas creen en lo que dicen. Sin embargo, va más allá de tal explicación. Según Hintikka “Moore basa su explicación de [una OM] en el hecho de que en la gran mayoría de los casos nosotros creemos aquello que decimos” (Hintikka 1962, p. 129). Pero señala más adelante: “No tengo por qué dar por supuesto que cuando se lleva a cabo un enunciado ( $q$  por ejemplo) bajo condiciones normales y en un tono de voz natural, haya que presumir que uno cree que  $q$  es verdadero. Para mis propósitos, basta con que siempre que uno diga que  $q$ , se suponga que uno puede creer aquello que uno dice (en el sentido de que  $q$  no ha de ser indefendible)” (Hintikka 1962, p. 131). Dice más: “Si estoy en lo cierto, quienquiera que diga [una OM] dice también algo que a la larga le resulta imposible de creer [...]. Esta imposibilidad no es sino una consecuencia de las propiedades lógicas de [una OM]” (Hintikka 1962, p. 129). Aunque Hintikka parezca guardar la intuición mooreana acerca del problema, va más allá cuando indica que la peculiaridad de una OM se entiende por medio del concepto de *indefendibilidad doxástica* (Hintikka 1962, p. 126). En este sentido, tal solución tendría importantes rasgos en común con el enfoque de Shoemaker, que se va a ver adelante, donde se defiende

Sin embargo, es posible encontrar virtudes en la contribución de Moore, aunque no solucione completamente el problema. Una cuestión importante que se deduce de su sugerencia, es la idea de que si alguien enuncia  $p$ , tal persona estaría, la mayoría de las veces, lista para corroborar tal aserción, en el sentido de comprometerse con  $p$ . Creo que tal enfoque permite hacer justicia a la intuición de que al aseverar  $p$  lo que se dice no sólo trata sobre el mundo, sino que también trata sobre uno mismo, al menos en el sentido de hacer explícito el propio punto de vista.

Cuando se presentó la paradoja, el hecho de haber una conjunción entre un enunciado supuestamente acerca del mundo y otro supuestamente acerca de la creencia de un sujeto, era lo que nos impedía encontrar una contradicción. Aunque Moore tampoco lo consiga, creo que su esfuerzo es el de aproximar el tópico de los dos enunciados. En el caso de Moore, él señala que las afirmaciones sobre el mundo no están separadas de la existencia de alguien que las afirma. Y si alguien las afirma, entendemos que se está comprometiendo con lo que dice.

#### *L. Wittgenstein*

La segunda propuesta de disolución de la paradoja se inspira en Wittgenstein<sup>11</sup>. Tal solución podría entenderse como la estrategia inversa a la primera, aunque por razones bastante distintas. Wittgenstein sugiere que decir “creo que  $p$ ” es solamente otra manera de decir “ $p$ ”. En sus palabras: “[E]l enunciado ‘Creo que va a llover’ tiene un sentido análogo a ‘Va a llover’, pero ‘Entonces creí que iba a llover’ no tiene un uso análogo a ‘Entonces llovió’” (Wittgenstein 1953, parte II, sección X).

Heal (1994, p. 20) entiende tal propuesta como la idea de que una persona que aprende un lenguaje está entrenada para usar, en ocasiones, “creo que  $p$ ” como sustituto de “ $p$ ”. De ese modo, la supuesta referencia a la creencia de la persona en  $p$  no añadiría nada, o no diferiría en nada de lo que ya se afirma con  $p$ . En este sentido, más que invertir la dirección, esta segunda solución se distingue de la primera por explicar dónde está la contradicción sin hacer referencia a lo que una persona implica con sus palabras. El verbo creer, usado en primera persona, podría entenderse aquí prácticamente como un término redundante<sup>12</sup>.

---

que el problema real estaría en creer en una OM. Pero, de nuevo, no se resumiría a tal posición porque Hintikka es también claro con respecto a tal posibilidad: “Yo no creo que tengan sentido únicamente aquellas formas verbales que puedan ser dichas y creídas (o conocidas) por el hablante. Lo que ocurre es que las formas verbales que no satisfacen esta condición resultan inútiles para la mayoría de los propósitos que uno espera que nuestro lenguaje cumpla” (Hintikka 1962, p. 131). Se agradece el comentario de un evaluador anónimo que indica que la propuesta que desarrollaré más adelante, que explica la peculiaridad de una OM en términos de principios de la racionalidad, podría compartir aspectos con tal propuesta de Hintikka, que toma “los principios de saber y creer” como explicación de su naturaleza absurda.

<sup>11</sup> Las propuestas de Heal (1994), Lee (2001) and Linville & Ring (1991), son ejemplos de variantes de la solución wittgensteiniana.

<sup>12</sup> “Si hubiera un verbo con el significado de ‘creer falsamente’, no tendría sentido usarlo en la primera persona del presente de indicativo” (Wittgenstein 1953, parte II, sección X).

Aplicando tal razonamiento a la paradoja, tenemos que una OM sería sencillamente una contradicción disfrazada entre “ $p$ ” y “no  $p$ ”<sup>13</sup>. En el caso de (2) es fácil reconocer tal resultado porque, si decir “creo que  $p$ ” es solamente otra alternativa a decir “ $p$ ”, afirmar (2) es afirmar ( $2^{**}$ ) “ $p$  y no  $p$ ”. Sin embargo, en el caso de (1) parece inútil dar solamente este paso ya que podemos usar “No creo que  $p$ ” para describir una situación en la que falta la creencia en  $p$  o en la que ignoramos si  $p$ . Así, afirmar (1) no nos lleva directamente a una contradicción entre  $p$  y no  $p$  porque “no creo que  $p$ ” no es solamente otra manera de decir “no  $p$ ”. Aquí, por lo tanto, la diferencia entre una OM del tipo (1) y del tipo (2) es todavía más importante, porque necesitamos condiciones complementarias para solucionar la primera. Esta es una de las críticas de peso al enfoque wittgensteiniano.

No obstante, hay quien interpreta las observaciones de Wittgenstein de una manera más amplia a la que estamos considerando, y que entiende que hay una doble dirección en la estrategia de mantener que “creo que  $p$ ” es solamente otra manera de decir “ $p$ ”. Lee (2001, p. 360-61) entiende que si Wittgenstein dice que “creo que  $p$ ” significa a grandes rasgos lo mismo que “ $p$ ”, la primera parte de la conjunción (1), “ $p$ ”, también puede entenderse como “creo que  $p$ ”. De tal manera, Lee sostiene que es posible hallar el absurdo en la forma de contradicción también en las oraciones del tipo (1).

La maniobra de Lee es interesante pero puede parecer demasiado *ad hoc* para establecer una salida consistente. Aceptar las dos posibilidades de sustitución (o bien “ $p$ ” por “creo que  $p$ ”, o bien “creo que  $p$ ” por “ $p$ ”) apela a principios de sustitución que solamente adquieren su sentido cuando interpretamos las palabras de alguien en busca de contradicciones, algo que, en general, no es lo que caracteriza a la interpretación. Es decir, la sustitución que permitiría encontrar una contradicción sólo emerge si ya tuviéramos como objetivo hallar la contradicción. Pero, en general, no se intenta buscar una contradicción en las palabras de alguien.

De cualquier manera, también quiero insistir en la importancia de la contribución wittgensteiniana a que expresiones como “creo que  $p$ ” no vayan separadas de expresiones como “ $p$ ”. Según Linville y Ring (1991, p. 302), que desarrollan un camino derivado del wittgensteiniano, la mayoría de los acercamientos partirían de la suposición de que la paradoja se establece porque estamos tratando una conjunción entre una oración auto-referencial (que dice algo sobre el hablante) y otra oración que tendría un contenido totalmente independiente de aquel. Al darnos cuenta que las dos partes de la conjunción tienen relación entre sí, somos capaces de percibir que sí hay un problema en una OM. En el caso específico de Wittgenstein, pienso que su posición revela la fina intuición que dice que cuando uno se pregunta a sí mismo si cree en  $p$ , lo que hace es mirar hacia el mundo en lugar de mirar hacia sí mismo.

Por ello, tal solución nos da espacio para que la entendamos como una prueba de que cuando alguien dice algo sobre sí mismo —como “creo que  $p$ ”— lo que tiene en mente es el mundo. En su famosa cita, Evans (1982, p. 225), dice que si alguien pregunta “¿Te parece que habrá una tercera guerra mundial?”, aquello en lo que pienso

---

<sup>13</sup> Linville y Ring (1991, p. 296) entienden el ejemplo de Wittgenstein “Creo que está lloviendo pero no está” como un absurdo, por consistir en dos enunciados contradictorios acerca del tiempo.

para contestar es lo mismo que pensaría si estuviera preguntando “¿habrá una tercera guerra mundial?”. Es decir, contesto a tal persona de la misma manera.

*S. Shoemaker*

La tercera vía para disolver la paradoja está inspirada por Shoemaker<sup>14</sup>. Tal alternativa puede llamarse “psicologismo sobre la paradoja de Moore” —como sugiere Kriegel (2004, p. 102)— y afirma que el carácter paradójico de las oraciones mooreanas se hereda de las “creencias mooreanas”. Las oraciones mooreanas son absurdas porque expresan creencias mooreanas (Kriegel 2004, p. 102).

Lo que sugiere este tipo de solución es que más problemático que afirmar una OM es creer en una. La paradoja no aparece solamente en el ámbito del discurso. De hecho, alguien podría defender que, en la práctica, muchas de las OM podrían incluso ganar sentido. De manera que lo que sí seguiría siendo problemático serían las OM en el nivel de las creencias. En cierta forma, es esto lo que Shoemaker (1995, pp. 75-76) defiende cuando denuncia la poca atención que se presta a la rareza de la idea de que alguien pudiera creer en un contenido proposicional con la forma de una OM. E insiste en que lo que realmente necesitaría explicarse sería por qué una persona no puede creer coherentemente que esté lloviendo y que no cree que está lloviendo.

Según Shoemaker (1995, p. 76) si una persona creyera en una OM inevitablemente estaría creyendo en una contradicción. Dice:

Considere la proposición que es la conjunción de esta proposición (Está lloviendo pero no creo que esté lloviendo) y la proposición de que el hablante cree tal proposición, o sea, la proposición expresada por la oración ‘Está lloviendo y no creo que está lloviendo, y que esto sea así es lo que creo’. Esto es auto-contradictorio. De este modo, es una característica de los contenidos de las oraciones mooreanas que si uno puede creer de hecho en ellas, el sujeto de tal creencia no podría creer que las tiene sin estar creyendo en una contradicción. (Shoemaker 1995, p. 76)

Shoemaker llega a este resultado por medio de algunas condiciones suplementarias, donde la principal es la caracterización de la relación entre creencias de primer y de segundo orden.

De acuerdo con Shoemaker (1995, p. 77) creer en algo lo compromete a uno a creer que cree en tal cosa, en el sentido de que si tal persona se pregunta si tiene tal creencia, deberá llegar a la conclusión que cree que cree. Es decir:

Si  $x$  cree que  $p$ ; entonces  $x$  está comprometida con la creencia de que ella misma cree que  $p$ .

$X$  está comprometida con la creencia de que  $p$ , si y solamente si, [...] si  $x$  toma en consideración si cree que  $p$ , entonces vendrá a creer que cree que  $p$ . (Kriegel 2004, p. 106)

De tal modo, la ocurrencia de una creencia mooreana implica en que si el sujeto se pregunta a sí mismo si cree en una proposición mooreana, esto debería llevarle a encontrarse en una contradicción.

Tal acercamiento entiende que la creencia de segundo orden es, de alguna manera, inherente a la de primer orden. Según la lectura de Kriegel (2004, p. 108), la creencia de segundo orden es una creencia disposicional y sería esto lo que le permitiría a Shoemaker construirla como ya presente, de alguna forma, en la de primer orden.

<sup>14</sup> Kriegel (2004) defiende una variante de tal solución.

Sin embargo, la crítica más inmediata a tal enfoque es que no es necesario que alguien se pregunte si tiene o no tiene determinada creencia para que sea un agente racional. En un caso en que una persona tuviera una creencia en una OM y no se preguntara sobre tal creencia, no habría impedimento alguno para que siguiera creyéndola. En este caso, tal persona no tendría en realidad la creencia de que cree en la proposición mooreana, y por lo tanto, no tendría de hecho (aunque potencialmente sí) creencias contradictorias (Kriegel 2004, p. 107).

Shoemaker defiende el principio de que el pensamiento constriñe el habla. En el caso de las OM, según tal posición, realmente no sería posible que alguien las dijera porque no sería posible creerlas, o por lo menos, no sería posible creerlas sin creer en una contradicción. Sin embargo, como Albritton (1995) bien observa, si alguien dice una OM no diríamos “no, no la crees porque no puedes creerla. Es imposible”. Sino que diríamos “¿qué?, ¿qué has dicho?”, indicando que es otro el problema involucrado en una OM.

Además de tales complicaciones, creo que una posición como la de Shoemaker, que soluciona la paradoja por medio de la relación entre creencias de primer y de segundo orden para indicar la contradicción entre creencias, tiene que comprometerse en algún momento con la tesis de la luminosidad. Es decir, en algún momento tiene que decir que el simple hecho de tener una creencia en  $p$  me pone en una situación en la que tengo una creencia sobre la creencia en  $p$ . Creo que los casos de fallo de accesibilidad de uno a su propia mente, como los casos de autoengaño, son indicadores de que hay algún problema con tal tesis<sup>15</sup>.

A pesar de estos problemas, una posición como la de Shoemaker conlleva también algunas intuiciones importantes y que no deberíamos depreciar. Una de ellas es que el problema con una OM no existe solamente en el acto de afirmarla, no es solo una transgresión del comportamiento. Es también y primariamente una transgresión acerca de lo que se afirma o se cree (Kriegel 2004, pp. 113-114). El problema no está solamente en que no existe una contradicción en mis palabras sino que en ellas están involucrados más aspectos de la vida mental del sujeto.

Las tres soluciones ofrecen importantes intuiciones que debemos tener en cuenta. Sin embargo, me gustaría ofrecer una cuarta, inspirada tanto por las soluciones que acabamos de ver, como por la manera davidsoniana de comprender el fenómeno de la comunicación. Mi propuesta sugiere que las diferentes soluciones pueden ser más útiles cuando se ven como complementarias que cuando lo hacen como opciones rivales. Y así, lo que voy a intentar es ofrecer una nueva explicación de por qué una OM es absurda.

### *3. La irracionalidad de una OM*

Una breve recapitulación: He sugerido que la primera manera de entender la paradoja expresa la intuición de que un enunciado sobre el mundo nos muestra algo sobre quién lo dijo. Nos muestra, por lo menos, qué tipo de visión tiene determinada perso-

---

<sup>15</sup> Es posible argumentar en contra de la tesis de la luminosidad como lo hace Williamson (2000), que ofrece una prueba de reducción al absurdo a partir de las premisas que la compondrían.

na, desde dónde ve las cosas y con qué se compromete. La segunda solución, he mantenido que no trataba solamente del camino inverso —la idea de que cuando alguien dice algo sobre sí mismo está diciendo algo sobre el mundo—, sino que, además, nos mostraba el carácter externo de la perspectiva de primera persona. Nos mostraba que cuestiones acerca de  $p$  y acerca de mi creencia en  $p$ , desde tal perspectiva, se podrían mezclar hasta el punto en que una fuera sencillamente una forma distinta de expresar la otra.

Me parece que ambas intuiciones son bastante sensatas y las voy a mantener en mi propuesta de solución respondiendo a la siguiente tesis:

T1: La distinción entre oraciones del tipo “ $p$ ” y del tipo “creo que  $p$ ” no corresponde a la diferencia entre aquellos enunciados que se refieren al mundo y los que se refieren a la vida psicológica de un sujeto, de manera que los haga enunciados independientes entre sí.

Esta tesis niega tanto que la forma general de las oraciones-tipo (sea esta “ $p$ ” o “creo que  $p$ ”) como su aseveración en actos de habla específicos determine de antemano que su contenido sea o bien acerca del mundo o bien acerca del sujeto. T1 no tiene que comprometerse con una postura más fuerte que redujera un tipo de enunciado al otro, como ocurrió con las dos primeras soluciones expuestas en la sección anterior. Es posible mantener que “ $p$ ” y “creo que  $p$ ” son enunciados distintos. Sin embargo, lo que T1 afirma es que no tiene sentido hablar en términos de qué referencia tiene cada tipo de frase de una manera apriorística; no hay razón para mantener que oraciones con la forma de “ $p$ ” tienen como referencia solamente al mundo y que oraciones con la forma de “creo que  $p$ ” tienen como referencia sólo la vida psicológica de un sujeto.

En este sentido, el papel de T1 es rechazar la presuposición de que una OM sea una conjunción de dos oraciones sobre dos materias distintas, lógicamente independientes entre sí, pero en conflicto. T1, además, no tiene la intención de reducir ninguna de las partes de una OM a la otra.

Con respecto a la tercera solución de la paradoja vista en la sección anterior voy a mantener una segunda tesis:

T2: lo absurdo de una oración mooreana se extiende tanto al ámbito del discurso, como al ámbito de la creencia.

T2, en realidad, no responde fielmente a la tercera propuesta porque aquella mantenía la tesis de la primacía de la paradoja en el nivel de las creencias sobre el nivel de las aserciones. Pero sí conserva la intuición de que el problema con una OM seguiría existiendo aunque no fuera dicha. La intención de T2 es dar cabida a tal intuición, pero guiada por la idea de que no hay una separación tajante entre el ámbito de la mente y el ámbito del discurso. Así, con referencia a la tercera propuesta, en lugar de mantener la idea de la primacía de la paradoja en el ámbito de las creencias, voy a mantener solamente que la paradoja se extiende a tal esfera.

Voy a desarrollar también una tercera tesis, pero antes intentaré dar crédito a la plausibilidad de T1 y de T2 en el sentido que he señalado. Para esto voy a utilizar las

consideraciones davidsonianas sobre nuestras interacciones lingüísticas, más específicamente usaré su idea de interpretación radical y las tesis involucradas en ella.

En líneas muy breves, la idea davidsoniana de interpretación radical nos invita a aceptar que la atribución de mentalidad a alguien y la atribución de significado a sus palabras ocurren simultáneamente. Es decir, intentar comprender lo que alguien dice y verlo como un agente que tiene una vida mental, no ocurre en momentos aislados ni distintos.

Pensemos en la famosa situación radical, en la que una intérprete se propone hacer un manual de traducción de un lenguaje de una sociedad completamente aislada de cualquier otra del planeta. Tal situación ya presupone que la intérprete sea capaz de reconocer la existencia de un comportamiento intencional y lingüístico en el interpretado. Ella no podría empezar a traducir cualquier sonido emitido por cualquier ser vivo con el que se encontrara. La identificación de la intencionalidad, o sea, la atribución de una mente a nuestro interpretado tiene que ser parte de la misma tarea de comprender sus palabras.

Si “Kurt profiere las palabras ‘Es regnet’ y bajo las condiciones correctas sabemos que está diciendo que está lloviendo”, lo hacemos porque al identificar su enunciación como lingüística e intencional, somos capaces de interpretar sus palabras: podemos decir qué significan, en ese momento, sus palabras (Davidson 1973, p. 125).

La propia actividad de describir comportamientos lingüísticos es ya una actividad de interpretación en la que, no solamente atribuimos determinados significados a las palabras de alguien, sino que también atribuimos estados mentales, como creencias y deseos. Traducir “Gavagai” sólo puede tener sentido como actividad si el intérprete entiende que tal sonido lo ha emitido un ser que se relaciona de determinada manera con el mundo, muy similar a la suya, y que quiere decir algo cuando dice “Gavagai”.

Es necesario notar que la visión de Davidson puede, y de hecho se extiende, desde la situación radical hasta nuestras situaciones más cotidianas. Según propone Davidson (1973, p. 125) toda comprensión del discurso del otro involucra una interpretación radical que podría entenderse, para hablantes de un mismo lenguaje, en términos de la pregunta: ¿cómo se puede determinar que el lenguaje es el mismo?

Si tal enfoque es plausible, parece que tanto T1 como T2 ganan fuerza. Alguien que niega T1 parece estar de acuerdo con que es posible oír “*p*” y no tener en cuenta nada de lo que representa que “*p*” sea dicho por alguien. En el caso de T2, alguien que la negara parece tener que comprometerse con la posibilidad de que hubiera algo en el ámbito de las creencias que no fuera accesible a nadie, o sea, que existieran cosas en la mente sin que existiera nada en el habla. Sin embargo, como he dicho, creo que tal contexto solamente es un estímulo para que uno acepte T1 y T2. No hay una consecuencia directa entre las posiciones davidsonianas y tales tesis, principalmente porque estamos hablando, en el caso de una OM, de oraciones particulares, mientras que Davidson defiende tesis generales para el lenguaje y la interpretación. Si tal enfoque es correcto, no hemos demostrado que “*p*” no refiera sólo al mundo ni “no creo que *p*” sólo al hablante, pero creo que nos anima a defenderlo. Y nos motiva a ser más holistas en ese respecto. Parece ser posible decir que tanto “*p*” como “creo que *p*” puedan hacer explícitos, a la vez, datos sobre el mundo y datos sobre el agente en el mundo.

Defender la plausibilidad de T1 y T2, sin embargo, tampoco explica qué equívoco hay en una OM. Para esto, antes de localizar donde se encuentra lo absurdo de una OM, expongo mi tercera tesis:

T3: Un enunciado es absurdo cuando no responde a los principios de la racionalidad<sup>16</sup>.

T3 llama la atención sobre otro presupuesto de la discusión tradicional sobre la paradoja de Moore y lo niega: que un enunciado sólo sería absurdo en presencia de una contradicción. Parece claro que un enunciado puede ser absurdo debido a muchas razones. No obstante, la paradoja se entiende tradicionalmente sólo en términos de la falta de una contradicción: “sabemos que una OM es absurda aunque no haya una contradicción presente en la oración en sí misma”<sup>17</sup>. La intención de T3, especialmente junto a T1, es dar cabida a una solución sin que tengamos que pasar por la reducción de una parte de la OM a otra para encontrar la oración contradictoria. Lo que T3 propone es que, resolver la paradoja no es indicar donde está la contradicción en la OM, sino explicar por qué una OM no es una aserción racional.

De vuelta al mismo panorama davidsoniano sobre la interpretación radical, otra tesis involucrada allí afirma que la actividad de interpretación solamente tiene éxito si se atribuye al interpretado un buen grado de consistencia y coherencia en sus comportamientos lingüísticos, además de un buen grado de acierto en sus enunciados. Esto hace parte de lo que Davidson llama “principio de caridad” que no es solamente la indicación de una estrategia para que interpretemos a alguien, sino la única manera que tenemos para interpretarnos mutuamente, sea esta una situación radical o cotidiana. Podemos entender, por lo tanto, que enunciados tomados como significativos e intencionales sólo pueden interpretarse si entendemos las palabras del interpretado sobre un fondo de coherencia y consistencia. Defenderé además de esto que, es necesario que supongamos que existe una unidad personal en nuestro interpretado. Llamaré a todo este conjunto de suposiciones que hacemos al interpretar a alguien “la suposición de la racionalidad” de nuestro interpretado.

Propongo que nos acordemos de la situación imaginaria del inicio del texto, en que mi amiga dice “está lloviendo pero no creo que esté lloviendo”. Como he dicho, estamos las dos frente a una ventana por la cual se puede ver claramente que llueve. Yo podría, por un lado, intentar reinterpretar sus palabras en otros términos, como, por ejemplo, una metáfora o como si quisiera decir otra cosa como “¡no creo que esté lloviendo! Se han estropeado mis planes de ir a la playa”<sup>18</sup>. Pero si no lo hago, tendría

---

<sup>16</sup> Los principios de la racionalidad a que T3 se refiere son más amplios que la noción de coherencia y consistencia lógica, como se mostrará más adelante.

<sup>17</sup> Williams (2006), por ejemplo, define la paradoja en estos términos.

<sup>18</sup> En este caso, si mi amiga quisiera decir esto, su enunciado no sería un caso genuino de OM. De hecho, existen dichos populares que tienen exactamente la misma forma que una OM, pero no lo son, porque son capaces de transmitir un sentido inteligible. Por ejemplo, el dicho “no creo en políticos honestos, pero haberlos haylos”, puede usarse perfectamente para mostrar la perplejidad frente a la existencia de tantos políticos corruptos. Otro uso común podría ser la oración “São Paulo es una ciudad muy desagradable, pero creo que no lo es”, para expresar la idea de que aunque la mayoría de la

que ver que mi amiga claramente no tiene un comportamiento racional. Como he sugerido, sé que se equivoca cuando dice que no cree que llueva, pero me deja sin espacio para convencerla de que está equivocada porque ha dicho al principio “llueve”. ¿Se ha equivocado o sabe que llueve? Parece que para entenderla, tendría que dividirla en dos personas, una que afirma la primera parte de su enunciado y otra que afirma la segunda parte. Sin embargo, no podría verla de tal manera, con tal división, sin pensar en ese momento que está siendo irracional.

En su trabajo de 1982, Davidson presenta un enfoque bastante interesante sobre la irracionalidad, y la toma como siendo un fenómeno que ocurre “dentro de la casa de la razón” y no ya fuera de la esfera de lo racional. De tal forma, la irracionalidad no se establecería para aquellas acciones de los demás que, bajo nuestro juicio, son poco razonables, sino para las creencias y acciones que resulten de un fallo dentro del ámbito mental de la propia persona. Según la definición que ofrece Davidson (1982, p. 179), serían casos de irracionalidad aquellos en los que hubiera una causa mental que no sería una razón para lo que causan, y es solamente en este sentido en el que podríamos decir que una persona es irracional cuando no está abierta a razones.

Pienso que este enfoque es bastante interesante para tratar de la irracionalidad porque por un lado, es más general que aquellos que entienden la irracionalidad como fallos en las reglas predeterminadas del razonamiento. Y por otro lado, la trata sin transformarla en otro fenómeno, es decir, consigue tratarla sin tomarla como un caso inexistente<sup>19</sup>.

Al definir casos de irracionalidad como casos de estados mentales que tienen como causa otros estados mentales que, sin embargo, no sirven de razones para los primeros, Davidson sostiene que si queremos explicar de hecho la irracionalidad, deberíamos asumir que la mente se divide en estructuras cuasi-independientes (Davidson, 1982, p. 181). Es solamente de ese modo como un estado mental podría causar otro sin jugar el papel de razón para él. Sostener tal idea no significa tener que enumerar, ni nombrar las diferentes partes de la mente. Tampoco necesitamos hablar de divisiones tajantes entre tales partes. Podemos mantenernos en un nivel más abstracto que este. Davidson señala incluso que términos como “partes de la mente”, o “partición de la mente” son términos engañosos si con ellos entendemos que lo que está en una región no puede estar en la otra.

Frente a tal acercamiento, he sugerido que una OM auténtica debería entenderse como un caso de irracionalidad porque hace evidente la división interna de la persona: en el ejemplo propuesto, mi amiga se encuentra en el mundo en parte considerando

---

gente la considere así, a mí no me parece que así sea. Es interesante que tales oraciones se parecen mucho a una OM, pero al utilizarlas expresando un sentido inteligible, no pueden entenderse como tal.

<sup>19</sup> Varios enfoques sobre la irracionalidad, según Davidson (1982), acaban por diluir el propio fenómeno. El principio de Platón, por ejemplo, que se basa en la doctrina de la racionalidad pura, diluiría la irracionalidad, transformándola en un caso más de ignorancia. En el extremo opuesto, según el principio de Medea, un acto irracional sería fruto de fuerzas ajenas a la persona. Según tal posición en un caso de acracia la persona actuaría en contra de su mejor juicio porque alguna fuerza ajena superaría a su voluntad, de lo que se sigue que los actos acráticos no serían intencionales.

que llueve y en parte considerando que no llueve<sup>20</sup>. Esto vuelve a mi amiga inconsistente, pero no a sus palabras.

En las condiciones de una interpretación radical, he mantenido que si queremos interpretar a alguien, debemos suponer la racionalidad de nuestro interpretado. Y suponer esto significa atribuir un fondo de verdad y de coherencia a sus palabras, y, además, la atribución de una unidad a su mente. En un cierto sentido, el “principio de caridad” davidsoniano es aquí ampliado porque la mera coherencia (del sistema de creencias, así como entre lo que se cree y lo que se dice), o incluso, la presunción de verdad en las palabras de alguien, no da plena cuenta de la racionalidad involucrada en la unidad personal. La explicación davidsoniana de la irracionalidad nos empuja a tomar la mente como compartimentada, pero esto no significa que tengamos que considerarla así en el momento de la interpretación. Al contrario, no podemos hacerlo. Según el propio Davidson, no hay dudas de que el principio de caridad exigido en una actividad de interpretación se opone a la idea de partición de la mente (1982, p. 184). Es decir, la irracionalidad se explica en términos de la posibilidad de la fragmentación de la mente, pero si de hecho no tomamos a nuestro interpretado como alguien irracional, no tenemos que considerarlo como fragmentado.

El problema con una OM podría formularse de la siguiente manera: en una situación de interpretación radical, sencillamente no podríamos empezar atribuyendo una OM a nuestro interpretado. Precisamente porque una OM es la expresión de un comportamiento irracional y, como vimos, tenemos motivos para aceptar que interpretar solamente es posible si asumimos que estamos frente a un ser racional. Está claro que tenemos comportamientos irracionales y es imprescindible que mantengamos un espacio en nuestras teorías para que quepan allí. Pero lo importante es que esto no puede ser la norma. En el caso de la lluvia, si estuviéramos tratando de un lenguaje total-

---

<sup>20</sup> Moran (2001, p. 69), por ejemplo, sugiere que casos de “acracia” o “autoengaño”, que son tradicionalmente tomados como fallos de la racionalidad, pueden entenderse en términos de la paradoja de Moore, justamente porque aparecen en situaciones de verdadera división con respecto a cómo uno se entiende a sí mismo: la división entre una actitud que tengo y otra que me atribuyo. Ofrece el siguiente paralelo: sé que lo que siento con respecto a él deberían ser celos (“creo que está lloviendo”), aunque no haya nada de lo que tener celos (pero no está lloviendo). Según lo expone, aquí tampoco tenemos una contradicción, porque tal enunciado representa un estado de cosas posible, pero sabemos que hay una gran tensión en tal caso.

Moran tiene un enfoque muy interesante sobre cómo tenemos acceso a nosotros mismos y una de sus tesis es que lo hacemos de dos modos: desde la perspectiva de primera persona y desde la perspectiva de tercera. Esto no significa que tengamos dos tipos de objetos distintos o partes específicas de lo mental accesibles a cada uno de los modos. Los dos modos de acceso son solamente distintas maneras de acceder a nosotros que nos posibilitan retractarnos o corregir nuestras propias creencias. Podría tener acceso a mí misma tanto por medio de uno como de otro, y corregirme si fuera el caso, justamente por la posibilidad de contrastar los dos puntos de vista sobre mí misma y mis comportamientos (Bensusan y Pinedo 2007). Comparado con la paradoja de Moore, un caso de “autoengaño”, por ejemplo, sería tan poco racional como un enunciado de una OM si el sujeto siguiera manteniendo dos puntos de vista contrapuestos sobre sí mismo. La posibilidad de que tengamos dos vías de acceso a nosotros mismos y que lleguemos a veces a resultados contrarios no nos da derecho a mantenerlos simultáneamente.

mente desconocido, entonces no podríamos empezar atribuyendo algo como una OM a las palabras de mi amiga.

Estamos obligados a tomar las palabras de alguien como teniendo que ver con su posición en el mundo y como manifestación de cómo esta persona en particular se relaciona con él. Pero además, buscamos que tal persona no se encuentre dividida en diversas partes que se relacionen de manera independiente unas de las otras, con los demás y con el mundo. Por esto, T3 tiene todavía más plausibilidad. Una OM es problemática, no porque haya una contradicción en enunciarla, sino porque nuestra idea de racionalidad supone la unidad personal. No se puede atribuir una OM bajo los principios de la racionalidad porque además de ir contra nuestras intuiciones lógicas también choca con nuestra concepción de nosotros mismos como personas. Y una de las ideas esenciales es que hay una unidad en cada uno de nosotros. Una idea, que a su vez, es totalmente distinta de la noción de accesibilidad total a todos nuestros estados mentales, y que tampoco es prueba de que no tengamos comportamientos contradictorios o tensiones dentro de nosotros mismos.

De forma que, creo que T1, T2 y T3 establecen un contexto desde donde podemos entender la peculiaridad de una OM con referencia a cómo estamos comprometidos con la idea de que actuamos lingüísticamente como seres racionales. Y en este caso, actuar lingüísticamente no puede suponer que las esferas del habla y de las creencias puedan estar separadas. Ni puede suponer que mantengo una posición neutra cuando hablo sobre mí o sobre el mundo; en ambas circunstancias aparecen rasgos sobre lo que pienso y sobre cómo veo el mundo. Si digo algo que solamente podría ser inteligible si me tomaran por dos personas, entonces es bastante probable que esté diciendo una OM.

## REFERENCIAS

- Albritton, R. (1995). "Comments on "Moore's Paradox and Self-knowledge""", *Philosophical Studies* 77, pp. 229-239.
- Baldwin, T. (1990). *G. E. Moore*. London and New York: Routledge.
- Bensusan, H., and M. Pinedo (2007). "When my Own Beliefs are not First-personal Enough", *Theoria* 58, pp. 35-41.
- Black, M. (1952). "Saying and Disbelieving", *Analysis* 13, pp. 25-33.
- Davidson, D. (1973). "Radical Interpretation", en *Inquiries into Truth and Interpretation*. Oxford: Clarendon Press, 1984, pp. 125-154.
- (1982). "Paradoxes of Irrationality", en *Problems of Rationality*. Oxford: Clarendon Press, 2004, pp. 169-188.
- Evans, G. (1982). *The Varieties of Reference*. New York: Oxford University Press.
- Heal, J. (1994). "Moore's Paradox: A Wittgensteinian Approach", *Mind* 103, pp. 5-24.
- Hintikka, J. (1962). *Saber y Creer*. (Traducción de Juan José Acero). Madrid: Editorial Tecnos, 1979.
- Kriegel, U. (2004). "Moore's Paradox and the Structure of Conscious Belief", *Erkenntnis* 61, pp. 99-121.
- Lec, B.D. (2001). "Moore's Paradox and Self-Ascribed Belief", *Erkenntnis* 55, pp. 359-370.
- Linville, L., and M. Ring (1991). "Moore's Paradox Revisited", *Synthese* 87, pp. 295-309.
- Moore, G.E. (1942). "A Reply to My Critics", en P. Schilpp (ed.), *The Philosophy of G.E. Moore*. La Salle: Open Court, pp. 535-677.
- (1944). "Russell's 'Theory of Descriptions'" en P. Schilpp (ed.), *The Philosophy of Bertrand Russell*. Evanston: Northwestern University, pp. 175-226.
- Moran, R. (2001). *Authority and Estrangement: An Essay on Self-Knowledge*. Princeton: Princeton University Press.

- Shoemaker, S. (1995). "Moore's Paradox and Self-Knowledge", en *The First-person Perspective and Other Essays*. Cambridge: Cambridge University Press, 1996, pp. 74-93.
- Williams, J.N. (1979). "Moore's Paradox – One or Two?", *Analysis* 39, pp. 141-142.
- (2006). "Wittgenstein, Moorean Absurdity and its Disappearance from Speech", *Synthese* 149, pp. 225-254.
- Williamson, T. (2000). *Knowledge and its Limits*. Oxford: Oxford University Press.
- Wittgenstein, L. (1953). *Investigaciones Filosóficas*. (Traducción de Alfonso García Suárez y Ulises Moulines). Barcelona: Crítica, 1988.

**Cristina BORGONI** tiene un Máster en Filosofía por la Universidad de Brasilia (Brasil) y es Investigadora FPI Contratada en Prácticas por la Universidad de Granada, en el Departamento de Filosofía I. Ha escrito diversos artículos sobre la filosofía de Donald Davidson y actualmente desarrolla su tesis doctoral sobre el externismo semántico y el autoconocimiento.

**DIRECCIÓN:** Departamento de Filosofía I, Edificio Facultad de Psicología, Universidad de Granada. Campus Cartuja, 18011 Granada, España. E-mail: cborgoni@ugr.es.