

Conocer la propia mente

Donald DAVIDSON *

(Traducción de Eva ALADRO)

(Abstracts y palabras clave al final del artículo)

Propuesto: 26 mayo 2005

Aceptado: 28 mayo 2005

El tipo de indicios de los que nos servimos para inferir aquello que los otros piensan no es ningún secreto: observamos sus acciones, leemos sus cartas, estudiamos sus expresiones, oímos lo que dicen, llegamos a saber sus historias y observamos sus relaciones con la sociedad. Cómo conseguimos juntar esa información en un retrato convincente de una mente es ya otro tema; sabemos cómo hacerlo, sin necesariamente saber lo que hacemos. A veces llego a saber aquello en lo que creo de la misma manera que cualquier otra persona lo llega a saber, observando cuanto digo y hago. Hay veces en que ésta es la única manera de tener acceso a mis propios pensamientos. Según Graham WALLAS,

* Donald DAVIDSON, fallecido en 2003, ha sido uno de los grandes filósofos del lenguaje y cognitivista del siglo XX. Heredero de la tradición filosófica centrada en el lenguaje, pero dentro de la línea cartesiana de disquisición lógica, se ha ocupado de todos los grandes problemas asociados al lenguaje y la comunicación, de los elementos constitutivos básicos de la cognición y la acción, y desde planteamientos siempre originales, como en su teoría interaccionista de la metáfora, o en sus estudios sobre creatividad e inteligencia. Aquí nos presenta la aportación al cognitivismo realizada por la filosofía y la reflexión deductiva en torno a temas como la autoridad de la primera persona en el uso del lenguaje y la cognición. Publicamos este texto gracias al permiso de su viuda, la profesora Marcia CAVELL.

la muchachita tenía alma de poeta y, cuando le decían que pensara bien lo que quería decir antes de hablar, ella respondía: ¿cómo puedo saber lo que pienso antes de decirlo?¹

Robert MOTHERWELL expresó un pensamiento semejante: “Yo diría que la mayor parte de los buenos pintores no saben lo que piensan hasta que lo pintan”.

Gilbert RYLE comprendía perfectamente al poeta y al pintor; afirmaba que necesariamente conocemos nuestras mentes del mismo modo como conocemos las mentes de los demás, observado cuanto decimos, hacemos y pintamos. RYLE estaba equivocado. Rara vez necesito de indicios o recurro a la observación para saber en lo que creo; normalmente, sé aquello que pienso antes de hablar o actuar. Incluso cuando tengo indicios, no suelo siquiera hacer uso de ellos. Puedo estar engañado acerca de mis propios pensamientos, por lo que recurrir a todo aquello que pueda ser públicamente determinado no es del todo irrelevante. Pero la posibilidad de poder engañarnos acerca de nuestros propios pensamientos no echa por tierra la suposición dominante de que una persona sabe en lo que cree; de un modo general, la creencia de que tenemos un pensamiento basta para justificarla. Pero aunque sea verdad y hasta algo obvio para la mayoría, el hecho, por lo que yo sé, no es fácilmente explicable. Aunque las estrategias que usamos para saber aquello que los otros piensan sean bastante obvias, por lo menos en términos generales, es un hecho bastante oscuro porque ocurre que, en nuestro propio caso, podemos con toda frecuencia saber aquello que pensamos sin necesitar indicios ni recurrir a la observación.

Dado que generalmente sabemos lo que pensamos (y deseamos, cuestionamos o pretendemos) sin necesitar, o servirnos, de indicios (incluso cuando están disponibles), nuestras declaraciones más sinceras respecto a nuestras actitudes presentes no están sujetas a los errores de las conclusiones basadas en indicios. Así, las pretensiones sinceras formuladas en primera persona del presente de indicativo relativas a pensamientos, aunque no sean infalibles ni incontestables, tienen una autoridad que ninguna pretensión en segunda o tercera persona, o en primera persona en otro tiempo verbal, podrían llegar a tener. Reconocer un hecho así no es, con todo, lo mismo que explicarlo.

Desde WITTGENSTEIN, se ha convertido en un lugar común minimizar nuestras dudas acerca del conocimiento que tenemos de las otras mentes,

¹ Graham WALLAS, *The Art of Thought*

diciendo que se trata de un aspecto esencial del uso que hacemos de ciertos predicados mentales, el hecho de aplicarlos a los otros en base a indicios comportamentales mientras que, por el contrario, los aplicamos a nosotros mismos sin recurrir a una ayuda tal. Esta observación es verdadera y, si está correctamente formulada, debería responder a alguien que se interrogue sobre el modo como podemos conocer las mentes ajenas. Pero como respuesta a un escéptico, la intuición de WITTGENSTEIN (si es que es de él) no debe proporcionar gran satisfacción. Puesto que, en primer lugar, es extraño pensar que las pretensiones hechas sin recurrir a los indicios o al comportamiento deben verse favorecidas en detrimento de aquellas que hacen uso de tales recursos. Es claro que, si no se puede recurrir a indicios para justificar una pretensión, entonces no podrá ser contestada poniendo en duda la verdad o la relevancia de los indicios. Pero tal cosa no basta para suponer que de un modo general las pretensiones que no se apoyan en indicios sean más fidedignas que aquellas que lo hacen. La segunda y principal dificultad es ésta. Es costumbre decir que aquello que cuenta como indicio para la aplicación de un concepto ayuda a definir el concepto o, por lo menos, pone restricciones a su identificación. Si dos conceptos dependen normalmente de diferentes criterios o tipos de indicios para su aplicación, deben ser conceptos diferentes. Así, si aquello que aparentemente es una misma expresión, se emplea unas veces de un modo correcto basándose en cierto tipo de indicios y otras veces se emplea con base a otro tipo de indicios (o a ninguno), la conclusión obvia sería la de que la expresión es ambigua. ¿Por qué entonces deberíamos pensar que un predicado como “x cree que Ras Danhan es la montaña más alta de Etiopía”, unas veces enunciado en base a indicios comportamentales y otras veces no, no es ambiguo? Si es ambiguo, entonces no hay ninguna razón para pensar que tiene el mismo significado cuando lo aplicamos a nosotros mismos o cuando lo aplicamos a los otros. Si afirmáramos (como debiéramos) que el carácter necesariamente público e interpersonal del lenguaje garantiza que aplicamos frecuentemente estos predicados de forma correcta a los otros y que por eso sabemos frecuentemente de hecho aquello que *los otros piensan*, entonces tenemos que plantear la cuestión de las razones en que nos apoyamos para pensar que sabemos aquello que (en el mismo sentido) *pensamos*. El tipo de respuesta wittgensteiniana puede resolver el problema de las otras mentes, pero crea un problema correspondiente al conocimiento de nuestra propia mente. Con todo, la correspondencia no es completa. El problema original de las otras mentes motivó la cuestión de cómo podemos saber que los otros tienen siquiera una mente. El problema con el que ahora nos enfrentamos tiene que plantearse de la siguiente forma: sé lo que buscar cuando atribuyo pensamientos a los otros. Usando criterios bastante diferentes (o ningunos), aplico los mismos predicados a mí mismo; entonces surge la duda escéptica de por qué debería yo pensar que son *pensamientos* aquello

que me atribuyo a mí mismo. Pero visto que los indicios de que me sirvo en el caso de los otros son públicos, no hay ninguna razón para no atribuir a mí mismo pensamientos de la misma manera que hago con los otros, a la manera de Graham WALLAS, Robert MOTHERWELL y Gilbert RYLE. En otras palabras, no lo hago, pero podría, tratar mis propios estados mentales del mismo modo en que trato los de los otros. Alguien que busque el mismo tipo de autoridad relativa a los pensamientos de los otros que la que tiene en relación con sus propios pensamientos se halla sin ninguna estrategia a la que recurrir. Así, la asimetría entre los casos continúa siendo un problema y es la autoridad de la primera persona la que causa el problema.

He sugerido una respuesta para este problema en otro artículo². En el mismo, defiendo que si prestáramos atención al modo como atribuimos pensamientos y significaciones a los otros, podríamos explicar la autoridad de la primera persona sin suscitar dudas escépticas. En los últimos años, de todos modos, algunos de los mismos hechos sobre la atribución de actitudes en los que yo me basaba para defender la autoridad de la primera persona han sido usados para atacar esa misma autoridad: se ha defendido, en base a aquello que se piensa que son nuevas pruebas, que aunque los métodos del intérprete en tercera persona determinen aquello que comúnmente pensamos que son los contenidos de la mente de un agente, los contenidos así determinados pueden ser desconocidos para el propio agente. En el presente artículo me detengo en algunos de estos argumentos y afirmo que no constituyen una verdadera amenaza a la autoridad de la primera persona. La explicación que dí en mi artículo sobre la asimetría entre atribuciones de actitudes a la primera persona y a las otras personas parece haberse fortalecido a la luz de las nuevas consideraciones, o de aquellas que de entre éstas parecen ser válidas.

Es adecuado precisar que el problema que aquí me preocupa no exige que nuestras creencias sobre nuestros propias actitudes presentes sean infalibles o incontestables. Podemos cometer errores y de hecho así lo hacemos, respecto a lo que pensamos, deseamos, aprobamos y pretendemos; siempre hay también la posibilidad de engañarnos a nosotros mismos. Pero esos casos, aunque no infrecuentes, no son ni podrán ser paradigmáticos. No trataré ahora esto, pero lo veo como uno de los hechos que deben ser explicados.

Dejando, pues, de lado el caso en el que nos engañamos a nosotros mismos y otros fenómenos anómalos o ambiguos, la cuestión es si podemos, sin ser irracionales, incoherentes o confusos, de modo simple y directo, pensar que tenemos cierta creencia que no tenemos o pensar que no tenemos una

² Donald DAVIDSON, "First person Authority", *Dialectica*, 38, (1984), pp. 101-111.

creencia que de hecho tenemos. Bastantes filósofos y psicólogos de perspectiva filosófica vienen defendiendo en los últimos tiempos posiciones que implican o sugieren que tal cosa puede suceder con la mayor de las facilidades –que debe, de hecho, ocurrir constantemente.

La amenaza ya se encontraba presente en la idea russelliana de las proposiciones que podían ser consideradas como verdaderas incluso cuando tenían “ingredientes” con los cuales la mente del conocedor no estaba familiarizada; a medida que el estudio de las actitudes *de re* (relativas a una cosa) se desarrolló, aumentó el peligro.

Pero fue Hillary PUTNAM quien dio el golpe de gracia. Pensemos en el argumento del que PUTNAM se sirvió en 1975 para mostrar que los significados, según él, pura y simplemente no están en la cabeza³. PUTNAM defiende de forma convincente que aquello que las palabras significan depende de algo más que de aquello que está en la cabeza de quien las piensa. Presenta algunas historias cuya moraleja es que las características de una historia y sus detalles de cómo alguien aprendió a usar una palabra influyen necesariamente su significado. De aquí parece seguirse que dos personas pueden estar en estados físicamente idénticos e incluso así decir cosas diferentes con las mismas palabras.

Las consecuencias son importantísimas. Pues si las personas pueden (habitualmente) expresar sus pensamientos correctamente mediante palabras, entonces sus pensamientos –sus creencias, deseos, intenciones, esperanzas, expectativas- también tienen que ser en parte identificados por acontecimientos y objetos exteriores a la persona. Si los significados no están en la cabeza, entonces, por lo que parece, tampoco lo están las creencias, ni los deseos, ni el resto.

Dado que algunos de ustedes podrían estar ya un poco cansados del sósias de PUTNAM en la Tierra Gemela, dejen que les cuente mi propia historia de ficción científica –si es que lo es-. Mi historia evita algunas dificultades irrelevantes de la historia de PUTNAM, aunque introduzca algunos problemas nuevos.⁴ (Pero más adelante volveré a la Tierra y a la Tierra Gemela).

³ Hilary PUTNAM, “The Meaning of “meaning”, reeditado en *Philosophical Papers*, vol II: *Mind, Language and Reality*, Cambridge University Press, 1975, p. 227.

⁴ No tengo ninguna pretensión de ser en esto original: Steven STICH usó un ejemplo muy parecido en “Autonomous Psychology and the Belief-Desire Thesis”, *The Monist*, 61 (1978), p. 573 ss. Debo matizar que no estoy sugiriendo que un objeto creado accidental o artificialmente no pueda pensar; el Hombre de los Pantanos necesita simplemente tiempo para adquirir una historia causal que

Imaginemos que un rayo cae sobre un árbol muerto en un pantano; yo me encuentro junto al árbol. Mi cuerpo se ve reducido a sus elementos básicos, en tanto, completamente por coincidencia (y a partir de moléculas diferentes) el árbol se transforma en una réplica mía. Mi réplica, el Hombre de los Pantanos, se mueve exactamente como yo hacía; siguiendo su naturaleza sale del pantano, encuentra y parece reconocer a mis amigos y parece responder a sus saludos en inglés. Va a mi casa y aparentemente escribe artículos sobre interpretación crítica. Nadie nota la diferencia.

Pero *hay* una diferencia. Mi réplica no puede reconocer a mis amigos; no puede *reconocer* nada, pues nunca conoció nada anteriormente. No puede saber el nombre de mis amigos (aunque ciertamente aparente saberlo), no puede acordarse de mi casa. No puede querer decir lo mismo que yo con la palabra “casa”, por ejemplo, pues el sonido “casa” que produce no fue aprendido en un contexto que le daría el significado correcto —o un significado cualquiera—. En verdad, no consigo ver cómo es que se puede decir que mi réplica puede querer decir lo que sea con los sonidos que hace, ni cómo es que puede tener pensamientos.

PUTNAM puede no estar de acuerdo con esta posición, pues él dice que si dos personas (u objetos) se encuentran en estados físicos significativamente semejantes, es absurdo pensar que sus estados psicológicos sean algo diferentes⁵. Sería un error afirmar con certeza que PUTNAM y yo discordamos en relación con este punto, empero, puesto que todavía no se ha esclarecido el modo como la expresión “estado psicológico” se usa.

PUTNAM piensa que muchos filósofos han asumido equivocadamente que estados psicológicos como creer y saber el significado de una palabra son simultáneamente 1) “internos” en el sentido en que no presuponen la existencia de otros individuos más allá del sujeto a quien se atribuye el estado y 2) que son éstos los estados propios que identificamos e individualizamos como hacemos con creencias y otras actitudes proposicionales. Dado que generalmente identificamos e individualizamos estados mentales y significados en términos parcialmente de relaciones con objetos y sucesos más allá del sujeto, PUTNAM cree que 1) y 2) caen por tierra: según él, no hay ningún estado que pueda satisfacer las dos condiciones.

pueda dar sentido a la pretensión de que él habla, se acuerda, se indentifica o piensa sobre las cosas del mundo (Volveré a este asunto

⁵ PUTNAM, H., op. cit., p. 144

PUTNAM denomina “limitados” (*narrow*) a los estados psicológicos que satisfacen la condición 1). Ve esos estados como solipsistas y los asocia a la perspectiva cartesiana de la mente. PUTNAM puede considerar estos estados como los únicos estados “verdaderos”; en la mayor parte de su artículo él omite el predicado “limitado”, a pesar del hecho de que los estados psicológicos limitados (así llamados) no corresponden a las actitudes proposicionales tal y como son normalmente identificadas. Ni toda la gente está convencida de que hay que hacer una distinción clara entre estados psicológicos limitados (o internos, cartesianos o individualistas- todos estos términos son moneda corriente) y los estados psicológicos identificados (si es que algunos lo son) en términos de hechos externos (sociales u otros). Por eso John SEARLE defendió que nuestras actitudes proposicionales comunes satisfacen la condición 1), no habiendo, por tanto, necesidad de estados que satisfagan la condición 2), en tanto que Tyler BURGE negó que existan, en sentido significativo alguno, actitudes proposicionales que satisfagan la condición 1)⁶. Pero parece haber unanimidad relativa al hecho de que ningún estado satisface ambas condiciones.

La tesis de este artículo es que no hay razón para pensar que estados mentales comunes no puedan satisfacer ambas condiciones 1) y 2): pienso que tales estados son “internos”, en la medida en que son idénticos a estados del cuerpo, siendo por tanto identificables sin referencia a objetos o sucesos exteriores al cuerpo; al mismo tiempo, son “no-individualistas” en sentido en que pueden ser, y generalmente lo son, identificados en parte por sus relaciones causales con sucesos e objetos exteriores al sujeto de quien son estados. Un corolario de esta tesis será que, contrariamente a lo que se piensa con frecuencia, la autoridad de la primera persona puede aplicarse sin contradicción a estados que son habitualmente identificados por sus relaciones con sucesos y objetos exteriores a la persona.

Comienzo por el corolario. ¿Por qué sería natural suponer que los estados que satisfacen la condición 2) no pueden ser conocidos por la persona que se encuentra en esos estados?

Ahora debo hablar de la Tierra Gemela de PUTNAM. Éste nos pide que imaginemos dos personas exactamente iguales físicamente y (por ello) iguales en lo que respecta a todos los estados psicológicos “limitados”. Una de las personas, un habitante de la Tierra, aprendió a usar la palabra “agua” cuando le mostraban agua, al leer u oír hablar de ella, etc. La otra, un habitante de la

⁶ Vid. John SEARLE, *Intentionality*, Cambridge University Press, 1983, y Tyler BURGE, “Individualism and Psychology”, *The Philosophical Review*, 95 (1986), pp. 3-45.

Tierra Gemela, aprendió a usar la palabra “agua” en circunstancias no diferentes en términos observacionales, pero la sustancia a la que fue expuesto no es agua sino una sustancia parecida, a la que podemos llamar “bagua”. En estas circunstancias, defiende PUTNAM, el primer hablante se refiere al agua cuando usa la palabra “agua”; su gemelo se refiere a la bagua cuando *él* usa la palabra “agua”. Parece, así, que estamos ante un caso en el que los estados psicológicos “limitados” son idénticos y, sin embargo, los hablantes quieren decir cosas diferentes al usar la misma palabra.

¿Y en cuanto respecta a los pensamientos de estos dos hablantes? El primero se dice a sí mismo, cuando ve un vaso de agua, “esto es un vaso de agua”; el segundo murmura exactamente los mismos sonidos para sí cuando ve un vaso de bagua. Ambos dicen la verdad, dado que sus palabras significan cosas diferentes. Y dado que ambos son sinceros, podemos pensar que creen en cosas diferentes: el primero cree que ve un vaso de agua ante sí y el segundo cree que hay un vaso de bagua delante *suya*. Pero ¿sabrán ellos aquello en lo que creen? Si los significados de sus palabras, y por lo tanto, las creencias expresadas en el uso de esas palabras, están parcialmente determinadas por factores externos acerca de los cuales los agentes nada saben, sus creencias y significados no están limitados en el sentido de PUTNAM. Por tanto, no hay nada en base a lo cual cualquiera de los dos hablantes pueda decir en qué estado se encuentra, pues no hay pistas internas ni externas en relación a la diferencia existente. Por lo que parece, deberíamos concluir que ninguno de los dos hablantes sabe aquello que dice o piensa. Esta conclusión fue la que sacaron explícitamente bastantes filósofos, entre los cuales se encuentra PUTNAM. PUTNAM afirma rechazar completamente la idea de que si existe una diferencia de significado, *debe* haber una diferencia en nuestros conceptos (o en nuestro estado psicológico). Lo que determina el significado y la extensión no es, por regla general, completamente conocido por el hablante⁷. Aquí “estado psicológico” quiere decir estado psicológico *limitado* y se parte del principio de que solamente estos estados son completamente conocidos. Jerry FODOR cree que las actitudes proposicionales comunes están (casi literalmente) en la cabeza, pero concuerda con PUTNAM en que *si* las actitudes proposicionales fuesen en parte identificadas por factores exteriores al agente, no estarían en la cabeza, y no serían necesariamente conocidas por el agente⁸. También John SEARLE, aunque por diferentes razones a las de FODOR, defiende que los significados están en nuestras

⁷ Cf. PUTNAM, H., “The Meaning of Meaning”, pp. 164-5.

⁸ Cf. Jerry FODOR, “Cognitive Science and the Twin Earth Problem”, *Notre Dame Journal of Formal Logic*, 23 (1982), p. 103. Vid. tb. del mismo autor “Methodological Solipsism Considered as a Research Strategy in Cognitive Psychology”, *The Behavioral and Brain Sciences*, 3 (1980).

cabezas (según él, no hay otro sitio donde puedan estar), pero parece aceptar la inferencia de que en caso de no ser así, se perdería la autoridad de la primera persona.⁹ La que constituye la formulación más simple de esta posición es la que se da en la introducción de Andrew WOODFIELD a un libro de ensayos sobre los objetos del pensamiento. Refiriéndose a la idea de que los contenidos de la mente vienen frecuentemente determinados por hechos externos y tal vez desconocidos para la persona que posee esa mente, dice:

Visto que la relación externa no está determinada subjetivamente, el sujeto no tiene ninguna autoridad sobre ella. Una tercera persona puede muy bien estar en mejor posición para saber cuál es el objeto en el que el sujeto está pensando y, por tanto, estar en una posición mejor para saber cuál era ese pensamiento¹⁰

Aquellos que aceptan la tesis de que los contenidos de las actitudes proposicionales se identifican en parte en términos de factores externos, parecen tener un problema semejante al problema del escéptico que descubre que podemos estar completamente engañados acerca del mundo “exterior”. En el caso presente, se evita el escepticismo tradicional en relación a los sentidos partiendo del principio de que es el propio mundo el que determina más o menos correctamente los contenidos de los pensamientos acerca del mundo. (El hablante que piensa estar delante de agua tiene probablemente razón, pues oyó la palabra “agua” utilizada en un contexto acuático; el hablante que piensa en bagua tiene probablemente razón pues aprendió la palabra en un contexto bacuático). Pero el escepticismo no se elimina; simplemente, se dirige hacia el conocimiento de nuestras propias mentes. Nuestras creencias habituales acerca del mundo exterior son (en esta perspectiva) dirigidas hacia el mundo, pero no sabemos aquello en lo que creemos.

Hay, es claro, una diferencia entre agua y bagua, y esa diferencia puede ser descubierta por medios normales, se quiera o no. Así, una persona puede descubrir aquello en lo que cree descubriendo la diferencia entre agua y bagua y descubriendo lo bastante acerca de sus propias relaciones con ambas, de forma que determine acerca de cuál de ellas habla él y tiene creencias. La conclusión escéptica a la que parece hemos llegado indica respecto a la extensión de la autoridad de la primera persona: es mucho más limitada de lo que pensábamos. Nuestras creencias acerca del mundo son en su mayoría verdaderas, pero podemos engañarnos fácilmente sobre lo que pensamos. Es una imagen inversa a la del escepticismo cartesiano.

⁹ SEARLE, J., *Intentionality*, cap. 8.

¹⁰ *Thought and Object*, Andrew WOODFIELD, ed., Clarendon Press, 1982, p. viii

Aquellos que defienden que los contenidos de nuestros pensamientos y los significados de nuestras palabras están frecuentemente fijados por factores de los cuales nada sabemos, no se han preocupado mucho con la consecuencia aparente de su perspectiva que aquí he resaltado. Ciertamente ellos percibieron que, si tuvieran razón, la idea cartesiana de que la única cosa de la que podemos tener certeza son los contenidos de nuestras propias mentes y la noción de Frege de los significados completamente “aprehendidos” tienen que ser erróneas. Pero no se esforzaron mucho, por lo que sé, en resolver el conflicto aparente entre sus perspectivas y la poderosa intuición de que la autoridad de la primera persona existe de hecho.

Una razón para esta falta de preocupación puede ser el hecho de que algunos de ellos parecen ver el problema como restringido a un tipo muy limitado de casos, casos en los que conceptos o palabras se asocian íntimamente a objetos que son seleccionados o referidos a través de nombres propios, de indexicales y palabras para especies naturales. Otros, sin embargo, defienden que los vínculos entre el lenguaje y el pensamiento por un lado y las cosas exteriores por otro están tan omnipresentes que ningún aspecto del pensamiento, tal como es habitualmente concebido, permanece intacto. En este sentido Daniel DENNETT afirma que debemos estar bien informados sobre el mundo en general e íntimamente ligados a él, a sus ocupantes y a sus propiedades, para que se pueda decir con justicia que tenemos creencias¹¹. Defiende que la identificación de *todas* las creencias está contaminada por factores externos y no-subjetivos que se sabe que funcionan en el tipo de caso que estamos tratando. BURGE también realiza la extensión de que nuestras creencias se ven afectadas por factores externos, aunque por razones que deja de explicar, aparentemente no ve eso como una amenaza a la autoridad de la primera persona.¹²

La cuestión tomó un rumbo inquietante. En el pasado, el behaviorismo se invocaba para mostrar cómo era posible que una persona supiese lo que pasaba en la cabeza de otra; el behaviorismo fue entonces rechazado en parte porque no podía explicar uno de los aspectos más obvios de los estados mentales: el hecho de que éstos son por regla general conocidos por la persona que los tiene sin recurrir a indicios comportamentales. La corriente más reciente, aunque no estrictamente behaviorista, identifica de nuevo los estados mentales parcialmente en términos de factores sociales y de otros

¹¹ Cf. DENNETT, D., “Beyond Belief”, en *Thought and Object*, p. 76.

¹² BURGE, Tyler, “Other Bodies”, en *Thought and Object*; “Individualism and the Mental”, en *Midwest Studies in Philosophy*, vol 4, Peter FRENCH, Theodor UEHLING, Howard WETTSEIN, eds., University of Minnesota Press, 1979; “Two Thought Experiments Reviewed”, *Notre Dame Journal of Formal Logic*, 23 (1982), pp. 284-93; “Individualism and Psychology”.

factores externos, permitiendo así, en esa medida, que puedan ser descubiertos públicamente. Pero reintroduce al mismo tiempo el problema de la justificación de la autoridad de la primera persona.

Aquellos que defienden la dimensión externa de los contenidos de los pensamientos tal como son habitualmente identificados e individualizados, reaccionaron de diversos modos. Una de las reacciones fue hacer una distinción entre los contenidos de la mente en cuanto determinados subjetiva e internamente, por un lado, y las creencias, los deseos y las intenciones habituales, tal como normalmente los atribuimos en base a relaciones sociales u otras relaciones externas, por otro lado. Éste es claramente el sentido del argumento de PUTNAM (aunque la palabra “agua” tenga diferentes significados y sea utilizada para expresar creencias diferentes cuando es usada para indicar agua y cuando lo es para indicar bagua, las personas que usan la palabra para estas finalidades diferentes pueden estar “en el mismo estado psicológico”). Jerry FODOR acepta la distinción para ciertos casos, pero defiende que la psicología debería asumir una posición de “solipsismo metodológico” (la expresión es de PUTNAM) o sea, debería tratar exclusivamente de estados internos, los estados psicológicos verdaderamente subjetivos que nada deben a sus relaciones con el mundo exterior.¹³

Steven STICH hace esencialmente la misma distinción, pero llega a una conclusión más severa: mientras FODOR piensa que nosotros sólo necesitamos entrar un poco en las actitudes proposicionales tal como éstas son habitualmente concebidas de modo que aislemos el elemento puramente subjetivo, STICH argumenta que los estados psicológicos tal como los vemos actualmente forman parte de una “psicología del sentido común” primitiva y confusa que debemos sustituir por una “ciencia cognitiva” todavía por inventar. El subtítulo de su último libro es “El argumento contra la creencia” (*The Case Against Belief*)¹⁴.

Sin duda aquellos que llegan a una conclusión tal se aseguraron de que el problema de la autoridad de la primera persona, por lo menos en la forma como yo lo he planteado, no puede resolverse. Pues el problema que yo planteé fue cómo explicar la asimetría entre el modo como una persona conoce sus estados mentales presentes y el modo como los otros los conocen. Los estados mentales en cuestión son creencias, deseos, intenciones, etc, tal como son habitualmente concebidos. Los que aceptan algo parecido a la

¹³ Jerry FODOR, “Methodological Solipsism Considered as a Research in Cognitive Psychology”

¹⁴ STICH, S., *From Folk Psychology to Cognitive Science*, MIT Press, 1983.

distinción de PUTNAM ni siquiera intentan explicar la autoridad de la primera persona en relación a estos estados: si hay una autoridad de la primera persona, debe estar asociada a otros estados bastante diferentes (en el caso de STICH, es difícil decir si podrá estar asociada a lo que sea).

Pienso que PUTNAM, BURGE, DENNETT, FODOR, STICH y otros tienen razón en llamar la atención sobre el hecho de que los estados mentales comunes, por lo menos las actitudes proposicionales, se identifican en parte por las relaciones con la sociedad y con el resto del ambiente, relaciones que pueden ser ignoradas por la persona que se encuentra en esos estados. También tienen razón, en mi opinión, en defender que por este motivo (que puede incluso ser el único) los conceptos de la “psicología del sentido común” no pueden verse integrados en un sistema de leyes coherente y omniabarcante del tipo creado por la física. Estos conceptos forman parte de una teoría del sentido común que describe, interpreta y explica el comportamiento humano de forma un tanto caótica pero que es indispensable (en mi opinión). Puedo imaginar una ciencia que tenga por objeto a las personas y que haya sido expurgada de “psicología del sentido común”, pero no consigo ver qué interés tendría esa ciencia. Éste no es, con todo, el tema de este artículo.

Lo que aquí me interesa es el descubrimiento intrigante de que aparentemente no sabemos lo que pensamos –por lo menos del modo como pensamos saberlo-. Esto es un verdadero enigma para aquellos que, como yo, creen que es verdad que los factores externos determinan en parte los contenidos de los pensamientos y que creen también que por regla general sabemos de hecho, y de una manera que los otros no saben, aquello en lo que pensamos. El problema surge porque la aceptación de los papeles identificadores e individualizadores de los factores externos lleva aparentemente a la conclusión de que nuestros pensamientos no pueden ser conocidos por nosotros.

¿Pero es válida esa conclusión? La respuesta depende, según creo, del modo como cada uno piense que la identificación de contenidos mentales depende de factores externos. Esta conclusión es válida, por ejemplo, para cualquier teoría que defienda que las actitudes proposicionales se identifican mediante objetos (tales como proposiciones, especímenes de proposiciones o representaciones) que no se encuentran en la mente o “antes” que ella y que contienen o incorporan (como “ingredientes”) objetos o sucesos exteriores al agente; es claro que nadie conoce las características infinitas de todos y cada uno de los objetos exteriores. Que la conclusión es válida según estas

suposiciones es algo generalmente aceptado.¹⁵ Con todo, por razones que comentaré más adelante, rechazo las suposiciones en las que la conclusión se apoya en este caso.

Tyler BURGE sugería que hay otro modo por el cual factores externos entran en la determinación de los contenidos del habla y del pensamiento. Una de sus “experiencias de pensamiento” me hace a mí callar. Hasta hace poco tiempo yo pensaba que la artritis era una inflamación de las articulaciones causada por depósitos de calcio; no sabía que cualquier inflamación de las articulaciones, por ejemplo la gota, era también artritis. Por eso cuando el médico me dijo (equivocadamente, como después se verificó) que yo tenía gota, pensé que tenía gota pero no pensé que tenía artritis. En este punto BURGE nos pide que imaginemos un mundo en el que yo fuera el mismo físicamente, pero en el que la palabra “artritis” solo se utilizase de hecho para inflamaciones de las articulaciones causadas por depósitos de calcio. De este modo, la frase “La gota no es una forma de artritis” hubiera sido verdadera, no falsa, y la creencia que yo habría expresado con esta frase no hubiera sido una creencia falsa de que la gota no es una forma de artritis, sino una creencia verdadera acerca de otra dolencia que no es artritis. Con todo, en el mundo imaginado todos mis estados físicos, mis “experiencias cualitativas internas”, mi comportamiento y mis disposiciones comportamentales son las mismas que en este mundo. Una *creencia* mía habría cambiado, pero yo no tendría ninguna razón para suponer que tal cosa hubiese sucedido, y por tanto, no se podría decir que yo supiese aquello en lo que creía.

BURGE realza el hecho de que su argumento depende de la

posibilidad de que alguien tenga una actitud proposicional a pesar de no dominar completamente una u otra noción en su contenido(...) si se quiere que la experiencia del pensamiento funcione, tenemos que encontrar en un momento dado que el sujeto cree en un contenido (o que tiene una actitud caracterizada por éste), a pesar de una comprensión incompleta o de un error de aplicación.¹⁶

De aquí parece seguirse que, si BURGE tiene razón, siempre que alguien está equivocado, confuso o parcialmente mal informado acerca del significado de una palabra, estará equivocado, confuso o parcialmente mal informado acerca de absolutamente todas sus creencias expresadas (o que pudieran serlo) con el uso de esa palabra. Dado que una tal “comprensión parcial” es “común

¹⁵ Vid. p. ej. EVANS, Gareth, *The Varieties of Reference*. Oxford University Press, 1982, pp. 45, 199, 201.

¹⁶ Tyler BURGE, “Individualism and the Mental”, p. 83.

o hasta incluso normal en un número bastante grande de expresiones en nuestro vocabulario”, según BURGE, debe ser de igual modo común o normal que estemos equivocados sobre lo que creemos (y por supuesto, acerca de aquello que recelamos, esperamos, deseamos que fuese, dudamos, etc).

BURGE aparentemente acepta esta conclusión; por lo menos es así como yo interpreto su objeción de que “la comprensión total de un contenido es de un modo general una condición necesaria para creer en el contenido”. Él rechaza explícitamente “el viejo modelo de acuerdo con el cual una persona tiene que conocer directamente o tiene que aprehender inmediatamente los contenidos de sus pensamientos (...). El *contenido* del pensamiento de una persona no está fijado por lo que sucede en su interior o por aquello a lo que tiene acceso simplemente por la vía de una reflexión cuidadosa”¹⁷

No sé muy bien cómo debo entender estas afirmaciones en la medida en que no sé si debo tomar en serio esta expresión de “conocer directamente” y de “aprehender inmediatamente” un contenido. Pero de cualquier modo estoy seguro de que si aquello que queremos decir y pensamos está determinado por los hábitos lingüísticos de nuestros prójimos del modo como postula Burge, entonces la autoridad de la primera persona queda seriamente en peligro. Dado que el grado y el carácter de este peligro me parecen incompatibles con aquello que sabemos acerca del tipo de conocimiento que tenemos de nuestras propias mentes, tengo que rechazar alguna de las premisas de BURGE. Concuero en que aquello que quiero decir y pienso no está “fijado” (exclusivamente) por aquello que sucede en mí, por tanto, lo que tengo que rechazar es la teoría de BURGE de cómo los factores sociales y otros externos controlan los contenidos de la mente de una persona.

Por varias razones, me siento conducido a minimizar la importancia de las características de nuestras atribuciones de actitudes a las que BURGE quiere traer nuestra atención. Supongamos que yo, que pienso que la palabra “artritis” se refiere a una inflamación de las articulaciones solamente si ésta viene causada por un depósito de calcio, y mi amigo Arturo, que sabe más que yo sobre el tema, decimos a Silvio con toda sinceridad que “Carlos tiene artritis”. Según BURGE, si nuestras circunstancias son más o menos las mismas, (ambos hablamos español con fluidez, ambos aplicamos frecuentemente la palabra “artritis” a casos verdaderos de artritis, etc.) entonces nuestras palabras en esta ocasión significan lo mismo, Arturo y yo queremos decir la misma cosa con nuestras palabras y expresamos la misma creencia. Mi error en lo que respecta al significado establecido de la palabra

¹⁷ Ib., pp. 90, 102, 104.

(o respecto a lo que es artritis) no crea diferencia alguna respecto de lo que quise decir o pensé en esa ocasión. Las pruebas que BURGE presenta para defender esta idea parecen basarse en su convicción de que esto es lo que alguien (no corrompido por la filosofía) diría acerca de mí y de Arturo. Dudo que BURGE tenga razón en este aspecto pero, incluso si tiene razón, no creo que eso concuerde con su idea. Las atribuciones de significados y actitudes habituales se basan en suposiciones abundantes y vagas sobre aquello que es y no es compartido (en términos lingüísticos así como en otros términos), por aquél que hace la atribución, la persona a quien se hace dicha atribución y la audiencia vista por quien hace la atribución. Cuando se verifica que algunas de estas suposiciones son falsas, podemos alterar las palabras que usamos para hacer la descripción, muchas veces de forma sustancial. Si la cuestión no es muy importante, tenemos tendencia a escoger el camino más fácil: tomamos las palabras de otro literalmente, aunque ello no refleje a la perfección algún aspecto del pensamiento de quien habla o de aquello que quiere decir. Pero esto no es porque nos veamos obligados (por lo menos fuera de un tribunal) a ser legalistas en relación a ello. Frecuentemente no lo somos. Si Silvio (no corrompido por la filosofía) dice a otra persona (tal vez a un médico que se encuentra lejos de allí y que intenta hacer un diagnóstico basándose en una descripción por teléfono) que Arturo y yo decimos y creemos que Carlos tiene artritis, él puede muy bien inducir a error al médico que *lo oye*. Si un peligro tal fuese probable, Silvio, más atento a los hechos, no diría simplemente: “Arturo y DAVIDSON creen los dos que Carlos tiene artritis”; sino que añadiría algo como, por ejemplo: “Pero DAVIDSON piensa que la artritis tiene que ser causada por depósitos de calcio”. La necesidad de una elucidación así muestra bien, en mi opinión, que la mera atribución no era completamente correcta; había una diferencia significativa en los pensamientos que Arturo y yo expresábamos al decir “Carlos tiene artritis”. Es claro que BURGE no está obligado a aceptar este argumento, dado que puede insistir en que la descripción es literalmente correcta pero puede, como puede ocurrir con cualquier descripción, inducir a error. Por otro lado, pienso que esta respuesta no tendría en cuenta el grado en el que los contenidos de una creencia dependen necesariamente de los contenidos de las otras creencias. Los pensamientos no son átomos aislados; por tanto, no puede haber reglas simples y rígidas para la atribución correcta de un pensamiento¹⁸.

¹⁸ BURGE sugiere que la razón por la cual normalmente pensamos que una persona quiere decir con sus palabras aquello que los demás en su comunidad lingüística quieren decir, sepa el hablante o no aquello que los otros quieren decir, es que “las personas son frecuentemente juzgadas, y se juzgan a sí mismas, de acuerdo con los patrones o modelos de la comunidad cuando se trata de una mala aplicación o de una mala comprensión”. Afirma también que tales casos “dependen de una cierta responsabilidad en la práctica comunitaria” (CF. “Individualism and the Mental”, p. 90). No pongo en duda el fenómeno, sino simplemente su importancia respecto a aquello que quiere mostrar.

Aunque yo rechace la insistencia de BURGE en que estamos obligados a dar a las palabras de una persona el significado que tienen en su comunidad lingüística, y a interpretar sus actitudes proposicionales basándonos en ese principio, pienso que hay otro sentido algo diferente pero muy importante en el cual los factores sociales controlan lo que un hablante puede querer decir con sus palabras. Si un hablante quiere ser comprendido, tiene que querer que sus palabras sean interpretadas de una cierta manera y tiene, por tanto, que querer proporcionar a su audiencia las pistas que necesite para llegar a la interpretación deseada. Esto es válido tanto en el caso de que el oyente domine el uso de una lengua que el hablante sabe, como en el caso de que el oyente sea un aprendiz de una lengua materna. Y el requisito de la aprendibilidad, de la interpretabilidad, que proporciona el factor social irreductible, es el que muestra por qué una persona no puede querer decir lo que sea con palabras que no puedan ser correctamente descifradas por otra persona. (BURGE parece llegar a esta misma conclusión en un artículo posterior¹⁹).

Me gustaría volver ahora al ejemplo de la Tierra Gemela de PUTNAM, que no depende de la idea de que el uso lingüístico social dicta (en condiciones más o menos modelizadas) aquello que los hablantes quieren decir con sus palabras, ni, obviamente, cuáles son los estados psicológicos (limitados) que tiene. Estoy convencido, como dije, de que PUTNAM tiene razón; aquello que nuestras palabras significan está en parte fijado por las circunstancias en las que aprendemos y usamos las palabras. El único ejemplo que PUTNAM nos da (el del agua) tal vez no sea suficiente para resolver definitivamente este aspecto, pues es posible argumentar que “agua” no se aplica solamente a una substancia con la misma estructura molecular del agua, sino también a substancias suficientemente parecidas a ella en términos estructurales como

(a) Es generalmente correcto considerar a las personas responsables de saber aquello que sus palabras quieren decir; en esos casos podemos considerar que defendían posiciones que desconocían o que no pensaban defender. Esto no tiene nada que ver (directamente) con aquello que querían decir con sus palabras ni con aquello que pensaban. (b) Como buenos ciudadanos y buenos padres procuramos animar las prácticas que fomentan las oportunidades de comunicación; usar las palabras como pensamos que los otros las usan puede fomentar la comunicación. Esta idea (justificada o no) puede ayudar a explicar por qué algunas personas tienen tendencia a atribuir significados y creencias de un modo normativo; pretenden animar al conformismo. (c) Un hablante que quiere ser comprendido tiene que querer que sus palabras sean interpretadas (luego interpretables) de acuerdo con ciertas reglas; este deseo puede satisfacerse usando las palabras del mismo modo que los otros las usan (aunque frecuentemente no ocurra tal cosa). De modo semejante, un oyente que quiera comprender a un hablante tiene que querer interpretar sus palabras como el hablante quería que fuesen interpretadas (ya sea esa interpretación la del “modelo” o no). Estas intenciones recíprocas adquieren importancia moral en un número infinito de situaciones que no tienen ninguna relación necesaria con la determinación de lo que alguien tiene en la cabeza.

¹⁹ Ver, por ej., “Two Thought Experiments Reviewed”

para ser inodoras y potables, para servir para nadar y navegar, etc. (Sé bien que estas consideraciones, como muchas otras a lo largo de este artículo, pueden sugerir que no sé reconocer una designación rígida. No sé desde luego). La cuestión no depende de casos así de especiales ni de cómo los resolvemos o deberíamos resolverlos. La cuestión depende simplemente de cómo la unión básica de palabras y cosas, o entre pensamientos y cosas, se establece. Defiendo, junto a BURGE y PUTNAM, si los he comprendido bien, que esa unión se establece por interacciones causales entre las personas y partes y aspectos del mundo. Las disposiciones para reaccionar de modo diferencial a objetos y sucesos así establecidas son vitales para interpretar correctamente los pensamientos y el discurso de una persona. Si no fuese así, no tendríamos medio alguno de descubrir aquello que los otros piensan o lo que quieren decir con sus palabras. El principio es tan simple y obvio como esto: una frase que alguien se vea motivado (o inducido) a aceptar como verdadera por la presencia y sólo por la presencia de la luna deberá querer decir algo como “Mira allí la luna”; el pensamiento expresado deberá ser el de que puede verse la luna; el pensamiento motivado por la presencia y sólo por la presencia de la luna deberá ser el de que la luna es visible. Deberá ser éste, teniendo en cuenta los comprensibles errores, los relatos de segunda mano, etc. No se trata del hecho de que todas las palabras y frases estén así tan directamente condicionadas por aquello a que se refieren; podemos perfectamente aprender a usar la palabra “luna” sin haberla visto nunca. Lo que se defiende es que todo pensamiento y todo lenguaje deben tener un fundamento en estas uniones históricas directas y que dichas uniones condicionan la interpretación de pensamientos y discursos. Tal vez sea bueno matizar que los argumentos a favor de esta idea no se basan en intuiciones relativas a aquello que diríamos en caso de que ciertos enunciados hipotéticos fuesen verdaderos. No necesitamos ficción científica ni experiencias de pensamiento²⁰.

Concuerdo, pues, con PUTNAM y con BURGE, que

(...) el contenido intencional de las actitudes proposicionales comunes (...) no puede ser explicado en términos de estados o procesos físicos, fenoménicos, causales-funcionales, computacionales o sintácticos, o por procesos que se especifican de

²⁰ BURGE ya dio ejemplos de “experiencias de pensamiento” que ni siquiera involucran al lenguaje; una de estas experiencias le lleva a afirmar que alguien que hubiese crecido en un ambiente sin aluminio no podría tener “pensamientos de aluminio” (“Individualism and Psychology”, p. 5). BURGE no nos dice por qué piensa de este modo, pero no es de ningún modo evidente que necesite suposiciones hipotéticas para defender su idea. Sea como fuese, las nuevas experiencias de pensamiento parecen basarse en intuiciones bastante diferentes de aquellas invocadas en “Individualism and the Mental”; el papel de las normas sociales en las nuevas experiencias no está claro y los hábitos lingüísticos de la comunidad parecen ser irrelevantes. Es posible que aquí la posición de BURGE esté cercana a la mía.

forma no intencional y que son definidos exclusivamente por el individuo, separado de su medio físico y social²¹.

Falta responder a la cuestión de si este hecho representa una amenaza para la autoridad de la primera persona, como BURGE parece defender y PUTNAM y otros ciertamente defienden. Rechacé uno de los argumentos de BURGE que, si fuera verdadero, constituiría esa amenaza. Pero existe la posición descrita en este último párrafo y que yo defiendo, ya lo hagan otros también o no, pues pienso que esta dosis de “externalismo” es necesaria para explicar cómo el lenguaje puede ser aprendido y cómo las palabras y actitudes pueden ser identificadas por un intérprete.

¿ Por qué PUTNAM piensa que si la referencia de una palabra (a veces) es fijada por la historia de cómo fue adquirida esa palabra, un usuario de la palabra puede perder la autoridad de la primera persona? PUTNAM afirma (correctamente, desde mi punto de vista) que dos personas pueden ser en todos los aspectos físicos relevantes (químicos, fisiológicos, etc) la misma, y con todo querer decir cosas diferentes con sus palabras y tener actitudes proposicionales diferentes (tal como éstas son normalmente identificadas). Las diferencias se deben a diferencias ambientales acerca de las cuales los dos agentes pueden, en determinados aspectos, no saber nada. ¿Por qué en estas condiciones, deberíamos suponer que estos agentes pueden no saber aquello que quieren decir y piensan? Hablar con ellos no nos revelará fácilmente la razón. Como matizamos, cada uno de ellos, en presencia de un vaso de agua o de bagua, dicen con toda sinceridad “Esto es un vaso de agua”. Si se encuentran en su medio ambiente respectivo, cada uno de ellos tienen razón; si cambiaran de tierra, cada uno de ellos estaría equivocado. Si preguntáramos a cada uno de ellos lo que quieren decir con la palabra “agua”, darían ambos una respuesta correcta usando, obviamente, las mismas palabras. Si les preguntáramos en qué creen, nos darían la respuesta correcta. Estas respuestas son correctas pues, aunque idénticas verbalmente, deben ser interpretadas de modo diferente. Qué es lo que ellos no saben (el sentido habitual de la autoridad de la primera persona) acerca de sus propios estados? Como ya vimos, PUTNAM distingue los estados de los que hemos hablado de los estados psicológicos “limitados”, que no presuponen la existencia de cualquier otro individuo más allá del sujeto que se encuentra en ese estado. Podemos comenzar ahora a pensar por qué está PUTNAM interesado en los estados psicológicos limitados. Una parte de la respuesta es que, obviamente, estos estados son los que piensa que tienen la propiedad “cartesiana” de ser conocidos de un modo especial por la persona que se encuentra en ellos. (la

²¹ “Two Thought Experiments Reviewed”, p. 288.

otra parte de la respuesta tiene que ver con la construcción de una “psicología científica”; ése es tema que no tratamos aquí).

El raciocinio depende, según creo, de dos suposiciones en gran medida incuestionadas. Son:

- 1) Si un pensamiento es identificado por una relación con algo exterior a la cabeza, no se encuentra completamente dentro de ésta (no está en la cabeza).
- 2) Si un pensamiento no está completamente en la cabeza, no puede ser “aprehendido” por la mente del modo exigido por la autoridad de la primera persona.

Que éste es el razonamiento de PUTNAM, lo sugiere su tesis de que, si dos cabezas son la misma, los estados psicológicos limitados deben ser los mismos. De este modo, si pensáramos que dos personas son iguales molécula a molécula (en el sentido en que dos corbatas pueden ser “idénticas”; podemos creer, si así lo deseáramos, que cada una de las personas piensa los mismos pensamientos verbalizados, tiene los mismos datos sensoriales, las mismas disposiciones, etc.), entonces es absurdo pensar que un estado psicológico es diferente de otro. Se trata, obviamente, de estados psicológicos limitados, no de los que normalmente atribuimos, que no están en la cabeza.²²

No es fácil decir exactamente de qué modo los pensamientos verbalizados, los datos sensoriales y las disposiciones pueden ser idénticas sin darle la vuelta a las corbatas; démosle la vuelta, pues, a las corbatas. Pues la idea es ésta: los estados psicológicos limitados de dos personas son idénticos cuando sus estados físicos no pueden distinguirse uno de otro. No valdría la pena que pongamos esto en duda, visto que Putnam ha podido definir los estados psicológicos limitados; lo que yo quiero poner en cuestión es la suposición (1), que llevó a la conclusión de que las actitudes proposicionales habituales no están en la cabeza y que por ello la autoridad de la primera persona no se aplica en su caso.

Tiene que quedar bien claro que simplemente el hecho de que los significados se identifiquen en parte por relaciones con objetos exteriores a la cabeza no implica que los significados no estén en la cabeza. Pensar lo contrario sería tan erróneo como argumentar que dado que yo estoy quemado, y por ello se presupone la existencia del sol, se sigue que la quemadura solar mía no es una condición de mi piel. Mi piel quemada por el sol puede ser

²² Cf. “The Meaning of meaning”, p. 227.

indistinta respecto a la piel de otra persona que se quemó por otros medios (nuestras pieles pueden ser idénticas en el “sentido de las corbatas”); sin embargo, uno de nosotros está realmente quemado por el sol, no ocurriendo esto al otro. Con este argumento es suficiente para mostrar que una consideración de los factores externos que entran en nuestros modos habituales de identificar estados mentales, no pone en duda una teoría de la identidad entre lo mental y lo físico. Andrew WOODFIELD parece pensar lo contrario. Escribe:

Ningún estado de hecho acerca de un objeto exterior al cerebro de la persona puede ser idéntico a un estado de ese cerebro, pues ningún estado cerebral presupone la existencia de un objeto externo²³

Los estados y los acontecimientos individuales no presuponen *conceptualmente* lo que sea en sí mismos; algunas de sus *descripciones*, sin embargo, pueden hacerlo. Mi abuelo paterno no me presupone, pero si alguien puede ser descrito como mi abuelo paterno, varias personas además de mi abuelo, incluido yo mismo, tienen que existir.

BURGE puede muy bien estar cometiendo un error semejante en el siguiente pasaje:

(...) Ningún suceso de un pensamiento (...) podría tener un contenido diferente y ser un espécimen del mismo suceso (...) Pues (...) el suceso de pensamiento de una persona no es idéntico a ningún suceso de ella descrito por la fisiología, la biología, la química o la física. Pues, sea b cualquier suceso dado descrito en los términos de una de las ciencias físicas que tiene lugar en el sujeto en tanto éste piensa un pensamiento relevante. Sea “b” de tal modo que denote el mismo suceso físico que tenga lugar en el sujeto en nuestra situación hipotética (...) b no necesita ser afectado por sucesos hipotéticos (que no alteran los contenidos de suceso de pensamiento). Así (...) b (el suceso físico) no es idéntico al suceso de pensamiento del sujeto.²⁴

BURGE no reclama como suyo el establecimiento de la premisa de este argumento ni, por ello, de su conclusión. Pero defiende que el rechazo de la premisa es “intuitivamente muy implausible”. Prosigue afirmando que “las teorías materialistas de la identidad habituaron a la imaginación a ver el contenido de un acontecimiento mental como cambiante, en tanto el acontecimiento permanece inalterable. Pero si tales suposiciones son hechos posibles o mera fantasía filosófica, se trata ya de otra cuestión”. Es porque él

²³ WOODFIELD, A., en *Thought and Object*, p. viii.

²⁴ “Individualism and the Mental”, p. 111.

piensa que rechazar la premisa es extremadamente improbable, por lo que defiende que “las teorías materialistas de la identidad se vuelven” ellas mismas “implausibles por las experiencias de pensamiento no individualista”.²⁵

Yo acepto la premisa de BURGE; pienso que su rechazo no es tan implausible como absurdo. Si dos acontecimientos mentales tienen contenidos diferentes son ciertamente acontecimientos diferentes. Según creo, lo que los casos imaginados por BURGE y por PUTNAM muestran (y lo que el ejemplo del Hombre de los Pantanos muestra de forma más directa) es que las personas que sean idénticas en todos los aspectos físicos relevantes (o “idénticas” en el sentido de las corbatas) pueden diferir en relación a lo que quieren decir o pensar, del mismo modo que pueden diferir en su condición de abuelos o de estar quemados por el sol. Pero está claro que hay *algo* diferente en torno a ellas, hasta en el mismo mundo físico; sus historias causales son diferentes.

Concluyo que el mero hecho de que los estados mentales y los acontecimientos comunes se individualizan en términos de relaciones con el mundo exterior no tiende a poner en duda las teorías de la identidad mental-física como tales. En conjunción con algunas otras suposiciones (plausibles), el “externalismo” de ciertos estados y acontecimientos mentales puede ser usado, según creo, para poner en duda teorías de la identidad que caen dentro del materialismo de los tipos (*type-type identity theories*); pero si suscribe alguna cosa, son las teorías de la identidad que caen dentro del materialismo de los especímenes (*token-token identity theories*). (No veo ninguna buena razón para llamar “materialistas” a todas las teorías de la identidad; si algunos acontecimientos mentales son acontecimientos físicos, tal cosa no los vuelve más físicos que mentales. La identidad es una relación de simetría.)

Siendo esto así, PUTNAM y WOODFIELD no tienen razón al afirmar que es absurdo pensar que dos personas podrían ser físicamente idénticas (en el sentido de las “corbatas”) y sin embargo diferir respecto a sus estados psicológicos habituales. BURGE, a menos que esté dispuesto a hacer un uso mucho más efectivo del que hace de las suposiciones esencialistas, está equivocado al pensar que ha demostrado que todas las teorías de identidad son implausibles. Por tanto, estamos perfectamente dispuestos para pensar que dos personas pueden ser idénticas en todos los aspectos físicos relevantes aunque

²⁵ Cf. “Individualism and Psychology”, p. 15, nota 7. Cf. tb. “Individualism and the Mental”, p. 111.

difieran psicológicamente: ésta es verdaderamente la posición del “monismo anómalo” que defendí en otro texto²⁶.

Un obstáculo al conocimiento no basado en indicios de nuestras propias actitudes proposicionales lo hemos eliminado ya. Pues si las creencias comunes y las otras actitudes pueden estar “en la cabeza” incluso cuando se identifican como actitudes que son en parte en los términos de aquello que no está en nuestra cabeza, entonces la amenaza a la autoridad de la primera persona no puede venir simplemente del hecho de que los factores externos son relevantes para la identificación de las actitudes.

Pero queda todavía una dificultad aparente. Es verdad que mi quemadura solar, aunque descriptible como tal solamente con relación al sol, es idéntica a una condición de mi piel que puede (pienso yo) ser descrita sin referencia a tales factores “externos”. Aún así, si, como científico especialista en todas las ciencias físicas, yo sólo tengo acceso a mi piel y no a la historia de su condición, entonces por hipótesis no hay manera alguna de poder yo decir que estoy quemado por el sol. Tal vez, entonces, alguien tenga autoridad de primera persona en cuanto a los contenidos de su mente sólo en tanto esos contenidos pueden describirse o descubrirse sin referencia a factores externos. En la medida en que los contenidos son identificados en términos de factores externos, la autoridad de la primera persona necesariamente cae. Puedo decir al observar mi piel cuál es su condición privada o “limitada” (*narrow*), pero nada de lo que yo pueda llegar a saber en este dominio restringido me podrá decir que yo estoy quemado por el sol. La diferencia entre referirse a y pensar en el agua y referirse a y pensar en bagua es como la diferencia entre estar quemado por el sol y tener la piel exactamente en la misma condición debido a otra cosa cualquiera. La diferencia semántica reside en el mundo exterior, más allá del alcance del conocimiento subjetivo o sublunar. Este puede muy bien ser el argumento.

Esta analogía, entre la visión limitada del dermatólogo y la visión de túnel del ojo de la mente, es esencialmente imperfecta. Depende para funcionar de una perspectiva defectuosa de la mente, una perspectiva que aquellos que han atacado el carácter subjetivo de los estados psicológicos comunes compartían con aquellos a quienes atacan. Si conseguimos abandonar esta perspectiva, la autoridad de la primera persona dejará de verse como un problema; ciertamente, la autoridad de la primera persona pasará a depender de los

²⁶ “Mental Events”, en DAVIDSON, Donald, *Essays on Actions and Events*, Oxford University Press, 1982.

factores sociales y públicos que se pensaba destruyen esa autoridad y pasará a explicarse por esos mismos factores.

Hay una perspectiva de la mente que quedó de tal modo arraigada en nuestra tradición filosófica que es casi imposible escapar a su influencia, incluso cuando se reconocen y repudian sus peores defectos. En una versión rudimentaria muy habitual la cosa se presenta así: la mente es un teatro en el cual el yo consciente asiste a una representación (de sombras en la pared). Esta representación se compone de “apariencias”, de datos sensoriales, de *qualia*, de aquello que viene dado en la experiencia. Lo que aparece en el escenario no son los objetos habituales del mundo que el ojo exterior registra y que el corazón aprecia, sino sus supuestos representantes. Lo que sea que sabemos sobre el mundo exterior depende de aquello que conseguimos captar por medio de pistas internas.

La dificultad aparente desde el inicio relativa a la descripción de lo mental, es descubrir cómo es posible reconstruir un trayecto del interior hacia el exterior. Otra dificultad evidente, aunque tal vez menos reconocida, es saber decir dónde se encuentra el yo en la imagen. Pues el yo parece por un lado incluir al teatro, al escenario, a los actores y al público; por otro lado, aquello que se conoce y se registra pertenece exclusivamente al público. Este segundo problema podía ser muy bien reformulado como el problema de la localización de los objetos de la mente: ¿están *en* la mente o simplemente son vistos *por* ella?

No me interesan aquí esos tales objetos de la mente (hoy en gran medida rechazados) en tanto dados a los sentidos, sino sus primos enjuiciadores, los objetos presumibles de las actitudes proposicionales, ya se piense en ellos como proposiciones, especímenes de las proposiciones, representaciones o fragmentos de “mentales”. La idea principal que yo quiero atacar es que éstas son entidades que la mente puede “mantener”, “captar”, “tener presente”, o con las cuales puede estar “familiarizada”. (Estas metáforas tienen probablemente un valor pedagógico: los *voyeurs* sólo quieren tener representaciones presentes en el ojo de su mente, mientras que los más agresivos las intentan captar; los ingleses pueden estar solamente familiarizados con los contenidos de la mente, en tanto otros tipos más amigables los cultivarán).

Es fácil ver cómo el descubrimiento de que los factores externos intervienen en la individualización de los estados mentales perturba la imagen de la mente que acabo de describir. Pues si estar en un estado mental es para la mente estar en una determinada relación como aprehender un objeto,

entonces sea lo que fuere que nos ayude a determinar el objeto de que se trata, debe ser igualmente aprehendido si la mente quiere saber en qué estado se encuentra. Esto se vuelve especialmente claro si un objeto exterior fuera un “ingrediente” del objeto presente a la mente. Pero en cualquier caso la persona que se encuentra en un estado mental puede no saber en qué estado mental se encuentra.

Es en este punto donde el concepto de lo subjetivo –relacionado con un estado mental- parece desmoronarse. De un lado están los estados internos reales en relación con los cuales la mente sostiene su autoridad; por el otro lado están los estados habituales de creencia, deseo, intención y significación, que están contaminados por sus necesarias relaciones con el mundo social y público.

Como analogía, tenemos el problema del especialista en quemaduras solares que no consigue descubrir, al examinar la piel, si se trata de un caso de quemadura solar normal o meramente de una condición idéntica con otra causa cualquiera. Podemos resolver este problema distinguiendo entre quemadura de sol y quemadura de algo parecido al sol; quemado por algo como el sol es simplemente idéntico a quemado por el sol, excepto en el hecho de que el sol no está necesariamente involucrado. El especialista puede detectar un caso de quemadura por algo como el sol simplemente mirando ésta, pero no ocurre lo mismo con un caso de quemadura solar. Esta solución resulta porque las dolencias de la piel, al contrario que los objetos de la mente, no tienen que ser de tal modo que haya una persona especial que pueda decir al mirarlas que se trata de una determinada dolencia u otra.

La solución en el caso de los estados mentales es diferente y más simple; es librarnos de la metáfora de los objetos presentes a la mente. La mayoría de nosotros ya ha abandonado hace mucho tiempo la idea de las percepciones, los datos sensoriales y el flujo de la experiencia como cosas “dadas” a la mente; debemos tratar los objetos proposicionales de la misma forma. Es obvio que las personas tienen creencias, deseos, dudas, etc.; pero admitir esto no es sugerir que las creencias, deseos y dudas sean *entidades* presentes o interiores a la mente, o que estar en esos estados exija la existencia de objetos mentales correspondientes.

Esto ya fue afirmado antes, de formas más o menos agresivas, pero por razones diferentes. Los escrúpulos ontológicos, por ejemplo, no me preocupan absolutamente nada. Siempre tendremos la necesidad de una reserva infinita de objetos para ayudarnos a describir y a identificar actitudes como la creencia; no estoy sugiriendo en modo alguno que las frases que expresan

creencias y las frases que atribuyen las otras actitudes no sean de naturaleza relacional. Lo que sugiero es que los objetos con los cuales relacionamos las personas de modo que describen sus actitudes no necesitan de ningún modo ser objetos *psicológicos*, objetos que deban ser aprehendidos, conocidos o mantenidos, por la persona cuyas actitudes son descritas.

Este modo de pensar también nos es familiar; QUINE lo sigue cuando sugiere que podemos usar nuestras propias frases para mantenernos a la par de los pensamientos de personas que no conocen nuestra lengua. El interés de Quine es semántico y nada dice en este contexto acerca de los aspectos epistemológicos y psicológicos de las actitudes. Necesitamos reunir todos estos diversos abordajes. Las frases acerca de las actitudes son relacionales; por razones *semánticas* tiene consecuentemente que haber objetos con los cuales relacionar a quienes tienen actitudes. Pero tener una actitud no es tener una entidad delante de la cual se halle la mente; por razones *psicológicas* y *epistemológicas* convincentes deberíamos negar que existan objetos de la mente.

La raíz del problema es el dogma de que tener un pensamiento es tener un objeto ante la mente. PUTNAM y FODOR (así como muchos otros) distinguieron dos tipos de objetos, los que son verdaderamente interiores y por tanto se hallan “ante la mente” o “aprehendidos” por ella, y los que identifican al pensamiento de forma habitual. Estoy de acuerdo en que ningún objeto puede cumplir estas dos funciones. PUTNAM (y algunos otros filósofos que he mencionado) piensa que la dificultad llega del hecho de que no se puede esperar que un objeto parcialmente identificado en términos de relaciones externas coincida con un objeto ante la mente, pues la mente puede desconocer la relación externa. Tal vez las cosas sean realmente así. Pero de ahí no se sigue que podamos encontrar *otros* objetos que aseguren la coincidencia deseada. Pues si el objeto *no está* en relación con el mundo, nunca podremos aprehender nada sobre el mundo teniendo ese objeto ante la mente; y, por las mismas razones, sería imposible detectar un pensamiento tal en otra persona. Así, parece que aquello que está ante la mente no puede incluir sus relaciones exteriores –su semántica. Por otro lado, si el objeto *está* relacionado con el mundo, entonces no puede estar completamente “ante la mente” en el sentido relavarte. Con todo, a menos que un objeto *semántico* pueda estar ante la mente *en su aspecto semántico*, el pensamiento, concebido en los términos de unos objetos tales, no puede escapar al destino de los datos sensoriales.

La dificultad básica es simple: si tener un pensamiento es tener un objeto “ante la mente”, y la identidad del objeto determina aquello que el

pensamiento es, entonces es siempre posible que estemos engañados acerca de aquello que pensamos. Pues a menos que sepamos *todo* sobre el objeto, habrá siempre sentidos en los cuales no sabemos de qué objeto se trata. Muchas han sido las tentativas de encontrar una relación entre una persona y un objeto que se mantenga en todos los contextos, caso único en el que pudiéramos decir intuitivamente de la persona que sabe de qué objeto se trata. Pero ninguno de estos intentos dio resultado y pienso que la razón para ello es obvia. El único objeto que satisfaría los requisitos relacionados con el estar “ante la mente” y de ser de tal modo que determine el contenido de un pensamiento debería, como las impresiones de HUME, “ser aquello que parece y parecer aquello que es”. No existen tales objetos, públicos o privados, abstractos o concretos.

Los argumentos de BURGE, PUTNAM, DENNETT, FODOR, STICH, KAPLAN, EVANS y de muchos otros para mostrar que las proposiciones no pueden *simultáneamente* determinar los contenidos de nuestros pensamientos y ser subjetivamente aseguradas son, en mi opinión, otras tantas variaciones del argumento simple y general que acabo de dibujar. No son sólo las proposiciones las que no pueden desempeñar esa función; ningún objeto podrá hacerlo.

Cuando nos hayamos liberado de la suposición de que los pensamientos deben tener objetos misteriosos, podremos ver cómo es que el hecho de que los estados mentales tal como comúnmente los concebimos se identifican en parte por su historia, no sólo no ataca el carácter interno de esos estados o amenaza la autoridad de la primera persona, sino que al mismo tiempo abre la puerta a una explicación de la autoridad de la primera persona. La explicación llega de la consciencia de que aquello que las palabras de una persona quieren decir depende en los casos más básicos de dos tipos de objetos y de sucesos que hacen, sólo con que la persona piense, que se puedan aplicar las palabras; de modo semejante para aquello acerca de lo cual son los pensamientos de la persona. Un intérprete de las palabras y pensamientos de otra persona tiene que depender de información dispersa, de una buena preparación y de conjeturas imaginativas para conseguir comprenderla. El propio agente, con todo, no está en posición de pensar si está en general usando sus propias palabras para referirse a los objetos y acontecimientos apropiados, dado que para lo que él regularmente las aplica da a sus palabras el significado que tienen y a sus pensamientos los contenidos que tienen. Es obvio que, en determinados casos particulares, la persona se puede equivocar en relación a aquello en lo que cree acerca del mundo: lo que es imposible es que se pueda engañar durante la mayor parte del tiempo. La razón es obvia: a menos que haya una presuposición de que el hablante sabe aquello que dice, o sea, que usa correctamente su propia lengua, no habrá nada que un intérprete deba

interpretar. Para decirlo de otro modo, no podemos concebir algo como una persona que aplica mal sus propias palabras de modo regular. La autoridad de la primera persona, el carácter social del lenguaje y los determinantes externos del pensamiento y del significado volverán a estar todos juntos, como debieron estar, en cuanto abandonemos el mito de lo subjetivo, la idea de que los pensamientos exigen objetos mentales.

RESUMEN:

Se recapitula la teoría de la autoridad de los enunciados en primera persona que el cartesianismo estableció ligando las bases del estudio de la cognición con la filosofía y el razonamiento. Mostrando los límites que distintos autores han encontrado en esta tesis, el autor aboga de todos modos por la autoridad que el uso de la primera persona imprime a su lenguaje, también marcado por los factores sociales del lenguaje y otros elementos externos centrados en la cooperación y rasgos culturales de los hablantes.

Palabras clave: teoría de la autoridad de los enunciados en primera persona, tradición filosófica cartesiana, ciencia cognitiva, filosofía, argumentos externalistas, Putnam, tierras gemelas, interaccionismo, contexto cultural, factores sociales en el uso del lenguaje.

ABSTRACT:

Revision of the theory of authority of first person enunciates which Cartesians imposed linking cognition study and philosophy and reasoning. Showing the limitation that diverse authors exposed, Davidson restates authority in first person enunciates' use by language, also give the social factors in this use as well as other external elements such as cooperation and cultural features of speakers.

Key words: authority theory of first person enunciates, Cartesian tradition, cognitive science, externalism, Putnam, Twin Earths, interactionism, cultural context, social factors in language use.

RÉSUMÉ:

Révision de la théorie de l' autorité de la première personne que le cartésianisme établit liant l' étude de la cognition et la philosophie et raison pour toujours. En montrant les limitations que différents auteurs ont indiqué, Davidson considère l' autorité de la première personne comme acceptable mais en ajoutant les facteurs sociaux ainsi que d' autres éléments extérieurs comme la coopération entre les parlants et le contexte culturel.

Mots clé: théorie de l' autorité de la première personne, tradition cartésienne, science cognitive, externalisme, Putnam, terres gémelles, interactionisme, contexte culturel, facteurs sociaux dans l' use du langage.