

**Acerca del externismo intencional:
Un análisis de los argumentos en favor de la naturaleza
amplia de los contenidos mentales**

Autor: Enrique Palacios Gómez

Tutora: Cristina Corredor Lanas

Facultad de Filosofía y Letras

Universidad de Valladolid

Índice

Contenido

Introducción	1
1. Intencionalidad.....	2
2. El problema de la individualización de los contenidos representacionales.	6
3. Antecedentes del externismo intencional: el externismo semántico.....	18
3.1 Teoría tradicional del significado	18
3.2 Crítica de Putnam a la teoría tradicional	22
3.3 Experimentos mentales de las tierras gemelas	24
3.4 Teoría alternativa de Putnam	27
4. Burge y el anti-individualismo de los contenidos mentales.....	34
4.1 Experimentos mentales	35
4.2 La interpretación de Burge de los experimentos mentales	37
4.3 Diferencias entre los argumentos de Burge y Putnam	43
5. Consideraciones críticas acerca de los argumentos de Burge.....	47
6. Un posible desarrollo ulterior: un análisis de los argumentos contra la posibilidad de individuar el contenido mental estrecho.....	49
7. Conclusiones	53

Introducción

Desde Hegel, el pensamiento acerca de lo mental ha estado dividido en dos paradigmas contrapuestos. El paradigma que podríamos llamar internista y el paradigma externista. El paradigma internista se encuentra caracterizado por la figura de Descartes y en su núcleo está el tratar de dar cuenta de lo mental “mirando únicamente dentro de la mente del sujeto” aislándolo del entorno. Por el contrario, el paradigma hegeliano se ha caracterizado por destacar la importancia de las instituciones sociales y otros factores “exteriores” al sujeto en la caracterización del mundo mental.

Dentro de la tradición analítica, esta contraposición se ha hecho evidente a raíz de la aparición de una serie de argumentos dirigidos a cuestionar algunos de los presupuestos en que se basaban las teorías internistas tradicionales. Entre estos argumentos, los argumentos basados en ciertos experimentos mentales han jugado un papel fundamental. En este trabajo trataremos de analizar dichos argumentos con el objetivo de entender mejor el debate en que se encuentran y su papel dentro de este.

El trabajo se estructura en cinco apartados principales. En el primero de ellos supone una introducción a algunos de los términos y conceptos que vamos a utilizar, en el segundo de ellos trataremos de identificar la tesis o conjunto de tesis en que se focaliza la disputa entre las posiciones internistas y externistas. En el tercer apartado haremos un análisis de ciertos argumentos que, en el marco de la discusión sobre el significado han inspirado, posiciones externistas y han servido de modelo para extender dichas posiciones y argumentos al terreno de la mente, que analizaremos en el cuarto apartado basándonos principalmente en la contribución de los experimentos mentales de Burge. En el quinto apartado señalaremos algunos de los problemas ligados a dichos argumentos. Acabaremos el trabajo con una breve reflexión sobre como, a raíz de los problemas indicados en el quinto apartado, quizá debiera replantearse la estrategia de argumentación externista en el futuro.

1. Intencionalidad¹

¿Qué tenemos en la mente? La primera respuesta a esta pregunta es tan tajante como poco informativa: en la mente tenemos estados mentales. Ahora bien, ¿Qué son los estados mentales, qué tipos hay y cuál es su naturaleza? La filosofía mente ha distinguido tradicionalmente dos grandes grupos de estados mentales: los estados intencionales (actitudes intencionales, actitudes proposicionales...) y los estados de conciencia.

Muy vagamente podemos definir la intencionalidad como la propiedad de los estados mentales que permite representar estados de cosas. Tiene que ver por tanto con el carácter representacional de la mente y es por ello por lo que comúnmente suele atribuírsela propiedades semánticas. Los estados de conciencia, por otro lado, serían todos aquellos estados mentales relacionados con el carácter puramente cualitativo de la experiencia, lo que supone para el sujeto tener una mente o estar en determinado estado. Son los llamados *estados cualitativos de conciencia* o estados fenoménicos². Al ser el debate entre externismo e internismo algo centrado principalmente en la naturaleza de los estados intencionales, dejaremos provisionalmente de lado la cuestión de los estados fenoménicos y pasaremos a caracterizar con un poco más detalle las características de los primeros.

Si bien la palabra “intencionalidad” tiene su origen en la escolástica medieval (deriva de la palabra latina *intentio* que a su vez deriva del verbo *intendere*), no es hasta el siglo XIX cuando el concepto es rescatado por Franz Brentano pasando a formar parte central de las reflexiones en torno a la naturaleza de lo mental (en torno a la ontología de la mente).

1 Para el desarrollo de este apartado me he basado fundamentalmente en (Pierre, 2014)

2 Lo cierto es que algunos filósofos han distinguido otros tipos de conciencia (Ned Block por ejemplo). El debate en torno a la naturaleza de la conciencia es ciertamente muy complejo, no se limita ni mucho menos a la filosofía analítica, y es uno de los más activos a día de hoy en el terreno de la filosofía de la mente. Así mismo la relación entre conciencia e intencionalidad es muy discutida y varía mucho en función del enfoque teórico que se asuma en cada caso. Por ello, tómese todo lo que aquí se va a decir como una obligada simplificación. Para una discusión algo más detallada remito a: (Siewert, 2006)

Conviene aclarar que el término intencionalidad no debe confundirse con el significado ordinario en castellano de la palabra “intención” (dicha palabra indicaría más bien un tipo de actitud intencional como pueden ser las creencias, deseos, miedos...). Así, como la etimología latina indica (derivado del verbo latino *tendere*), la idea relevante aquí es la de direccionalidad. La intencionalidad tiene que ver por tanto, con la capacidad de la mente de apuntar hacia algo, de dirigirse hacia algo. Ese algo es lo que de ahora en adelante denominaremos contenido intencional u objeto intencional.

Franz Brentano anticipó en su libro *Psychology From an Empirical Standpoint* algunas de las características que ocuparán el centro de las discusiones filosóficas en torno a la naturaleza de la intencionalidad. Así, en dos famosos párrafos de dicha obra escribe:

Todo fenómeno psíquico está caracterizado por lo que los escolásticos de la Edad Media llamaron la inexistencia intencional (o mental) de un objeto, y que nosotros llamaríamos, si bien no con expresiones enteramente inequívocas, la referencia a un contenido, la dirección a un objeto (por el cual no hay que entender aquí una realidad) o la objetividad inmanente. Todo fenómeno psíquico contiene en sí algo como su objeto, aunque no todos del mismo modo. En la representación hay algo representado; en el juicio, algo admitido o rechazado; en el amor, amado; en el odio, odiado; en el apetito, apetecido, etc. (...)

Esta inexistencia intencional es característica, exclusivamente, de los fenómenos mentales. Ningún fenómeno físico exhibe nada semejante. Podemos, por tanto, definir los fenómenos mentales diciendo que son aquellos fenómenos que contienen un objeto intencionalmente en su interior.³

En los anteriores dos párrafos Brentano traza las líneas de lo que será todo un programa de investigación basado en tres tesis principales: i) es constitutivo del fenómeno de la intencionalidad el estar dirigido hacia cosas distintas a él mismo. Los verbos amar, creer, odiar, desear percibir, etc expresan todas ellas actitudes intencionales dirigidas hacia un objeto (Hay siempre un objeto un “algo” que se cree, ama, percibe...) ii) La segunda de las tesis se refiere a los objetos o contenidos intencionales. Es característico de ellos el

3 (Brentano, 1874, 88-89) Cita tomada de (Pierre, 2014)

exhibir lo que Brentano llama la “inexistencia intencional” iii) Brentano mantiene la controvertida tesis de que la intencional es la marca de lo mental. Todos los estados mentales y sólo los estados mentales exhiben intencionalidad.

Todas estas tesis, no exentas de ambigüedad, muestran un poco las líneas principales de controversia en la filosofía del siglo XX en torno a la naturaleza de la intencionalidad y lo mental. ¿Qué quiere decir eso de que los estados intencionales apunten hacia un objeto intencional, y qué es exactamente el objeto intencional? ¿Qué es eso de la “inexistencia intencional” de los objetos intencionales? ¿Quiere decir que los objetos intencionales hacia los que la mente se dirige sean internos a la mente misma (existen sólo en la mente, o quiere decir que podemos tener creencias, deseos.. sobre cosas que no existen, o ambas cosas a la vez? ¿Qué tiene de verdad la hipótesis de Brentano que reza que la intencionalidad sea lo que define lo mental? ¿dónde quedan entonces los estados de conciencia y como se relaciona la intencionalidad con ellos?

Si bien analizar la disputa en torno a estas cuestiones supondría exponer toda la historia de la filosofía de la mente, propondremos provisionalmente cuatro características principales que clásicamente se han asociado a la intencionalidad:

1) La capacidad de dirigirse hacia algo: Como hemos dicho, los estados intencionales se componen de dos elementos: i) lo que podríamos llamar la actitud intencional, la actitud que el sujeto tiene hacia el objeto o contenido intencional: creer, desear percibir..., y ii) el objeto o el contenido hacia el cual se dirige la actitud intencional: lo que se cree, desea, percibe...

2) El contenido intencional representa estados de cosas: tradicionalmente se ha considerado que el contenido intencional tiene propiedades semánticas y por tanto, al menos en el caso de actitudes proposicionales (creer, pretender...), en principio debería ser posible analizar sus condiciones de verdad. Las creencias pueden ser verdaderas o falsas, las pretensiones satisfechas o insatisfechas etc...

3) Podemos creer cosas distintas sobre el mismo objeto: En efecto, una persona puede creer que el agua quita la sed y no creer que el H₂O quita la sed al desconocer que el agua es siempre H₂O

4) Podemos creer, desear, esperar.. cosas acerca de objetos que no existen: Pedro puede creer que los ángeles tienen alas. En ese caso, si miramos el contenido intencional de la creencia de Pedro (los ángeles tienen alas), hay un objeto que no existe del cual se predica algo. En la misma línea está el hecho de que podamos negar la existencia de determinados objetos que no existen.

Podría considerarse que estos cuatro rasgos manifiestan el carácter perspectivista de los estados intencionales. Así, parece *a priori*, que el contenido intencional de los estados mentales manifiesta el modo en que el sujeto concibe la realidad, independientemente de que ésta sea o no de ese modo. Cuando, para analizar la intencionalidad, pasamos del nivel oscuro de la mente al nivel del lenguaje, estos rasgos se ponen de manifiesto a través de los contextos intensionales que se generan en el lenguaje con que describimos las actitudes proposicionales de un sujeto.

La intencionalidad del lenguaje se define como no extensionalidad. Dos propiedades son características de la extensionalidad:

i) Un contexto lingüístico es extensional, si dos términos correferenciales pueden sustituirse mutuamente *salva veritate*. Esto no ocurre en las descripciones de actitudes intencionales como puede verse en el siguiente ejemplo.

1. Ana cree que Fósforo brilla

2. Ana cree que Héspero brilla

En efecto, si bien los términos “Fósforo” y “Héspero” tienen la misma extensión, no siempre pueden intercambiarse libremente sin alterar el valor de verdad de la oración completa ya que, según lo que crea Ana, 1 puede ser verdadero y no 2 o viceversa.

b) En un lenguaje extensional se puede aplicar la regla de la generalización existencial de tal modo que si Pa es verdadero, también es verdadero que Existe algo que cumple P.

Esto no ocurre con las adscripciones⁴ de creencias, deseos etc.. . Del hecho de que Juan crea que los ángeles tienen alas no se sigue que existen esas cosas de las cuales Juan cree que tienen alas.

Resumanos todo lo que hemos dicho hasta ahora: el ser humano tiene estados mentales muy diversos. Podemos agrupar provisionalmente todos estos estados mentales en dos grandes grupos. En el primero tendríamos los estados de conciencia (un orgasmo, la sensación o experiencia interna que tengo ahora mismo de la luz del ordenador, la sensación cualitativa que me produce teclear las teclas...), y en el segundo los estados intencionales (creo que estoy escribiendo el trabajo de fin de grado, tengo la intención acabar pronto el trabajo, deseo no suspender...). Los estados intencionales tienen dos componentes: la actitud intencional en la que me encuentro (creencia, deseo...) y el contenido intencional al que se dirige mi actitud (lo que creo deseo...). Así mismo, los contenidos intencionales, los contenidos de lo que creo. deseo... tienen propiedades representacionales (Si amo a Laura, el contenido “Laura” puede tener una referencia, o si creo que estoy escribiendo, el contenido que representa la frase “estoy escribiendo” tiene condiciones de verdad). Sin embargo, puedo tener creencias (o actitudes intencionales en general) distintas sobre objetos que son el mismo (Puedo creer que Cicerón fue un gran orador y no creer que Marco Tulio lo fue) y puedo tener actitudes intencionales acerca de objetos que no existen (un niño, por ejemplo, podría creer que el mago Merín tenía poderes mágicos.). Estos dos últimos rasgos se manifiestan al nivel de lenguaje ya que las representaciones lingüísticas de estados intencionales generan contextos intensionales.

2. El problema de la individualización de los contenidos representacionales.⁵

Hasta ahora hemos hecho un repaso a algunas de las características que generalmente suelen asociarse a los estados intencionales y al mundo de lo mental. Tras esta introducción estamos ya en situación de comenzar a concretar el debate en que nos

4 Llamaremos adscripciones de creencias deseos etc... a la transcripción lingüística del contenido intencional. Por ejemplo, supongamos que Pedro está en el estado intencional de creer que la nieve es blanca. Dicha actitud puede transcribirse al lenguaje mediante la frase “Pedro cree que la nieve es blanca”. En dicha frase, el verbo “creer” representa la actitud intencional en la que se encuentra el sujeto (la actitud de creencia), y lo que viene tras la cláusula “que”, es decir, “la nieve es blanca” es la adscripción de creencia y representa el contenido intencional de la actitud intencional de Pedro.

5 Para el desarrollo de este apartado me he basado principalmente en (Burge, 1993), sobre todo. Aunque también hay elementos tomados de (Burge, 2007, Introduction) y (Lawrence, Margolis, 2002)

vamos a centrar a lo largo de este trabajo. Mi intención en este apartado es identificar el punto de discrepancia entre las posturas internistas y las externistas, un paso necesario si queremos entender bien el objetivo o la diana a la que se dirigen argumentos externistas que analizaremos en las páginas que siguen.

Para esta labor seguiremos el siguiente método: explicaremos las tesis o presupuestos en que se basa la teoría tradicional de la mente para, a continuación identificar en qué tesis o conjunto de tesis discreparían las teorías internistas y externistas de la mente. Es necesario aclarar que, en este sentido, hemos decidido sacrificar el rigor histórico en favor de la claridad. Téngase en cuenta por tanto, que todo lo que se va a decir en este apartado es una idealización de la teoría tradicional de la mente⁶.

El marco general del debate que vamos a tratar (el debate entre internismo y externismo) se encuadra dentro del debate general sobre la naturaleza de los estados intencionales ¿qué determina la naturaleza de los estados intencionales? Entendemos aquí “naturaleza” de un modo no excesivamente técnico y sin, en principio, asumir compromisos metafísicos muy fuertes.

Cuando nos preguntamos por la naturaleza de un estado intencional nos estamos preguntando por las condiciones necesarias (o necesarias y suficientes) que hacen que un estado intencional sea el estado intencional que es. En la medida en que podamos aproximarnos a dichas condiciones, ¿qué determina “constitutivamente” que un estado intencional sea uno y no otro? Dicho de otro modo: la dependencia constitutiva de una cosa respecto de otra afecta a la primera de modo que, si A depende constitutivamente de B, A no sería A sin esa relación con B. Es por ello por lo que la pregunta sobre la naturaleza de los estados intencionales se ha formulado en términos de la pregunta por la “individualización” de los estados intencionales⁷. El tipo de estado⁸ mental en que

6 De hecho pocos filósofos suscribirían punto por punto las tesis que vamos a presentar. La razón de presentar esta postura idealizada es que muchas veces se ha tomado como punto de partida de la discusión y para los propósitos que tenemos en este trabajo puede muy bien servir para darnos el tablero sobre el que se desarrolla el juego entre las posturas externistas e internistas.

7 Es decir, cuándo dos estados intencionales son el mismo y cuando son distintos. Qué condiciones delimitan, individualizan el estado intencional. Es decir, que condiciones tiene que cumplir un estado intencional para ser el estado intencional que es y no otro distinto.

8 Ejemplo: El deseo de Ana de comer pan es un tipo de estado intencional, la creencia de Juan de que los árboles tienen hojas es otro tipo de estado intencional, etc... En general siempre que podamos decir que dos estados intencionales son distintos, diremos que son dos “tipos” de estados intencionales

alguien se encuentre depende de ciertas características, ¿qué características son esas?

Como hemos dicho antes, los estados intencionales tienen dos componentes: la actitud intencional (creencia, deseo..) y el contenido intencional u objeto intencional (lo que se cree, desea...). Es generalmente reconocido que ambos componentes determinan la naturaleza del estado intencional; por extensión, es también generalmente reconocido en filosofía de la mente que un elemento esencial a la hora de decidir en qué estado intencional se encuentra alguien, es el contenido intencional de ese estado. Es más, en principio parece que siempre que el contenido intencional sea distinto, será distinto también el estado intencional. Podemos formular, por tanto la siguiente tesis.

Tesis 1: El tipo de contenido intencional es un elemento constitutivo el tipo de estado intencional (es decir, si cambia el contenido intencional, cambia el estado intencional)

Por este motivo nuestra discusión (la discusión entre el internismo intencional⁹ y el externismo intencional que vamos a tratar en este trabajo) se encuadrará dentro del debate sobre naturaleza de los contenidos intencionales (o representacionales, como también los llamaremos).

Deberemos entonces preguntarnos ahora por las condiciones, los factores que hacen que un contenido intencional sea el contenido intencional que es (el problema de la industrialización de los contenidos intencionales) ¿qué determina constitutivamente la naturaleza del contenido intencional? El debate entre el internismo y el externismo sobre los contenidos intencionales se encuadra dentro del debate en torno a la respuesta a esta pregunta.

Como hemos dicho, el contenido intencional tiene propiedades semánticas (en el caso de los contenidos proposicionales, estos tienen condiciones de verdad). Sin embargo, no sólo eso. Es generalmente admitido que dos contenidos intencionales no pueden ser el mismo contenido intencional si sus propiedades semánticas difieren, por tanto:

9 Esperamos que los términos “internismo intencional” y “externismo intencional” no lleven a equívoco. Con ellos nos estamos refiriendo a el internismo y el externismo acerca de los contenidos intencionales respectivamente. Los términos están sacados de (Moya, 2004, segunda parte, cap. 8 y 9)

Tesis 2. Las propiedades semánticas de los contenidos intencionales son un elemento constitutivo del tipo de contenido intencional (es decir, si dos contenidos intencionales tienen propiedades semánticas distintas, entonces dichos contenidos intencionales son distintos)

Tener en cuenta las propiedades semánticas parece entonces un paso (como mínimo) necesario a la hora de individualizar los contenidos intencionales.

También es generalmente reconocido que el contenido intencional está formado por conceptos (el pensamiento de que el vaso está encima de la mesa, está formado por los conceptos de vaso, mesa y estar encima de). Los conceptos, según la visión tradicional idealizada que estamos exponiendo, constituyen el modo en que la persona que piensa, piensa sobre objetos, propiedades, relaciones...(elementos que componen el contenido intencional). Además los conceptos tienen, según esta visión, propiedades semánticas. Dichas propiedades semánticas son constitutivas de los conceptos (en el sentido en que hemos visto hasta ahora). Así, el concepto de “eclipse”, es el concepto que es en parte porque (dado el mundo) se refiere o puede representar a todos los eclipses y a nada más. Es decir, si el concepto se pudiera aplicar sobre cosas distintas a los eclipses y sólo los eclipses, si representara otros objetos del mundo por ejemplo, entonces no sería el concepto de “eclipse”. Así mismo, en tanto que componentes del contenido intencional, los conceptos son elementos constitutivos del contenido intencional. Si el estado intencional estuviera compuesto por conceptos distintos, en principio, el contenido intencional resultante sería distinto. No sólo eso, en la teoría tradicional se asume que los conceptos (con sus propiedades semánticas) y su modo de articulación determinan el contenido intencional (y sus propiedades semánticas); ; en este sentido, en nuestra teoría idealizada estamos asumiendo la composicionalidad del contenido intencional.

Tesis 3. Las propiedades semánticas de los conceptos son constitutivas del tipo de conceptos.

Tesis 4. Los conceptos (en tanto que componentes de los contenidos intencionales) determinan, junto a su modo de articulación, los contenidos intencionales complejos.

Estos últimos rasgos de la teoría tradicional, hacen posible explicar el modo en que los

conceptos permiten al sujeto pensante relacionar distintos tipos de pensamientos entre sí de acuerdo con patrones inferenciales¹⁰. Así, consideramos que el pensamiento de que todos los perros son animales y el pensamiento de que Fido es un perro, comparten un concepto, lo que permite al sujeto pensante inferir, de acuerdo con las reglas deductivas, que Fido es un perro.

Las tesis que hemos ido enumerando hasta ahora indican dos cosas. Primero: la naturaleza de los estados intencionales dependerá en gran medida de la naturaleza de los contenidos intencionales, y ésta, a su vez, de la naturaleza de los conceptos mentales. Segundo: hemos hecho hincapié en las propiedades representacionales (semánticas) de los contenidos intencionales y los conceptos, así como su papel a la hora de distinguir unos conceptos o unos contenidos de otros. Ahora bien, desde la postura tradicional de la mente, es admitido que no son lo mismo los conceptos o contenidos mentales, de las entidades que pueden representar. Más bien la naturaleza de lo primero, determinaría lo segundo. Así, por ejemplo, el concepto que tengo de silla sería algo así como “el modo de presentación”, la idea de silla que tengo en la mente (sea lo que sea) que indica el conjunto de cosas que representa y es distinta a esas cosas que representa. Esto vale igualmente para los contenidos proposicionales. El contenido proposicional (un pensamiento completo) determinaría las condiciones de verdad, es decir, determina que tiene que pasar en el mundo para que dicho contenido sea verdadero pero es distinto a esos hechos del mundo.¹¹

Tesis 5. Los conceptos (los modos de pensamiento) son distintos de las entidades que representan.

Tesis 5.1 Del mismo modo, los contenidos proposicionales son distintos de el conjunto de circunstancias en que tales contenidos son verdaderos o falsos

Tesis 6. La naturaleza de los conceptos determina el conjunto de entidades que

10 Esto sólo es posible gracias a la composicionalidad del contenido intencional que hemos asumido.

11 El concepto de “gato” determina que objetos del mundo caen bajo su extensión, es decir, el conjunto de todos los gatos. Pero, obviamente, el concepto de “gato” es distinto de los gatos que danzan por ahí. Lo mismo ocurre con los contenidos proposicionales. El pensamiento de que “el gato está encima de la mesa” determina que tiene que pasar en el mundo para que ese pensamiento sea verdadero: tiene que pasar que efectivamente haya un gato encima de la mesa. Pero, igual que antes, el pensamiento es distinto a el hecho del mundo.

representa dicho concepto.

Tesis 6.2¹² Del mismo modo, la naturaleza del contenido intencional determina las condiciones en que podemos decir que tal contenido es verdadero o falso (en caso de contenidos proposicionales).

Si repasamos el recorrido que hemos hecho hasta ahora de las principales tesis que caracterizan la versión idealizada de la teoría tradicional de la mente que hemos expuesto, es fácil percatarse de que en último término todo depende de cuál sea la naturaleza última de los conceptos. Las condiciones necesarias y suficientes que hacen que un concepto sea el concepto que es, determinarán sus propiedades representacionales, y los conceptos y sus propiedades semánticas, determinarán la naturaleza y las propiedades semánticas de los contenidos representacionales complejos. Parece entonces un paso obligado plantearse la naturaleza última de los conceptos si queremos averiguar la naturaleza de los contenidos intencionales (marco en el que se encuadra el debate entre internismo y externismo acerca de los contenidos mentales).

Sobre este problema, el problema de la naturaleza de los conceptos (y por tanto de los contenidos intencionales), se han dado muchas respuestas a lo largo de la historia de la filosofía (dualismo, funcionalismo, conductismo, materialismo de la identidad...¹³), sin embargo, podemos agrupar todas estas soluciones en dos grandes grupos: Internismo y externismo. La confrontación entre estas dos posturas ha sido un tema muy vivo en la filosofía de la mente de los últimos años. Cabe decir, sin embargo, que definir ambas posturas resulta complicado dado que bajo las etiquetas “internismo” o “externismo” se encuentran concepciones de la mente muy distintas entre sí, muchas veces incluso contradictorias. En este sentido, el problema de definir etiquetas como “internismo” puede considerarse análogo al problema de definir etiquetas como “naturalismo”.

Por otro lado es necesario tener en cuenta que las etiquetas “internismo” y “externismo” surgieron muy recientemente y sólo a raíz de la identificación de una serie de posturas (las posturas que se han denominado posteriormente “externistas”) que de algún modo

12 Que no lleve a error la numeración. EL hecho de que hayamos puesto 5.1 y 6.2 no quiere decir que sean tesis que estén a otro nivel o algo por el estilo. Igualmente podríamos haber numerado: 5, 6, 7 y 8...

13 Ver más en (Burge, 1992) para un recorrido histórico sobre las principales teorías de la mente.

violentaban algunas de las intuiciones en las que descansaban las teorías clásicas de la mente. Podemos entonces afirmar que éste es un debate fruto de la confrontación que recientemente han hecho ciertas filosofías a las estrategias clásicas de individualización de los contenidos mentales, a los presupuestos que las teorías clásicas tienen sobre la naturaleza de los contenidos mentales (lo cual no quiere decir que no haya habido posturas externistas anteriormente) ¿qué presupuestos son esos?

Antes hemos mencionado dos aspectos esenciales acerca del contenido intencional: éste puede versar sobre objetos que no existen (podemos tener conceptos que si bien su naturaleza es ser representacional, tienen propiedades semánticas, no hay ningún objeto del mundo al que refieran), y además podemos tener contenidos mentales distintos que, individuados por su extensión, no difieren en nada: una persona puede creer que Héspero brilla y no creer que Fósforo brilla y, sin embargo, las condiciones de verdad de ambas sentencias en el lenguaje son las mismas y Héspero es el mismo objeto que Fósforo (el dilema de Frege). En definitiva, el carácter perspectivista que habíamos mencionado antes. Si añadimos a estas consideraciones la intuición fuerte de que tenemos acceso directo, conocimiento directo de nuestros estados mentales (lo que en ocasiones se ha denominado “la autoridad de la primera persona”), y por tanto somos capaces de distinguirlos entre sí “mirando hacia nosotros mismos”, es fácil derivar la tesis de que la naturaleza de nuestros conceptos (y por tanto de nuestros contenidos intencionales) no depende constitutivamente de elementos externos al sujeto. Sino únicamente de elementos internos al sujeto.

Tesis 7. Las condiciones necesarias y suficientes que hacen que un concepto sea el concepto que es son internas al sujeto

7a)¹⁴ 7 se ve respaldado por el hecho de que tenemos conocimiento directo y completo de nuestros conceptos

7b) 7 se ve respaldado por el hecho de que nuestra mente representa el mundo independientemente de cómo sea éste. Por ejemplo, podemos tener conceptos sobre cosas que no existen.

¹⁴ El motivo de escribir 7a y 7b (en lugar de 7.1, 7.2 tal y como hicimos antes) es que estas tesis no están al mismo nivel que 7. Son más bien tesis que respaldan o motivan defender 7.

Tesis 8. De 7 y 4 podemos concluir: las condiciones necesarias y suficientes que hacen que un contenido intencional sea el contenido intencional que es son internas al sujeto

La postura tradicional de la mente considera así, que los factores que determinan la naturaleza de todos los conceptos, y por tanto, todo el contenido mental son factores internos al propio sujeto, a su propia mente o cerebro, de tal modo que el mundo externo es indiferente, en el sentido en que no determina constitutivamente la naturaleza del contenido intencional. En definitiva, para saber qué piensa un sujeto hemos de mirar dentro de él, en su mente o cerebro¹⁵ y no fuera. Estas tesis de la teoría tradicional de la mente (7 y 8) son las definitorias del internismo acerca de los contenidos intencionales. Son internistas todas aquellas posturas que individualizan o definen los contenidos representacionales (tanto simples, en el caso de los conceptos, como complejos) en términos de características internas al sujeto (independientemente de cuál sea la postura ontológica concreta sobre la mente que subscribamos)¹⁶

Esta postura ha sido defendida por muchos autores a lo largo de la historia de la filosofía. Descartes y Hume, suelen considerarse representantes paradigmáticos de esta propuesta. El experimento mental del genio maligno cartesiano es un claro ejemplo de que, aun si el mundo externo fuera completamente distinto, yo seguiría teniendo los mismos contenidos mentales, por lo tanto estos solo dependen de factores internos al sujeto.

La postura internista también se adapta bien a propuestas materialistas. Stich¹⁷ lo ha defendido a través de el llamado “principio de autonomía psicológica”. Según dicho principio, un sujeto y una replica exacta de él, tendrían los mismos contenidos mentales aún si el mundo fuera completamente distinto. Lo cual no sólo parece compatible con 7,

15 No trataremos en este trabajo las diversas posturas que hay sobre la relación entre mente y cerebro. Para una explisición algo más detallada del tema, remito al lector a (Moya, 2004, segunda parte)

16 Es necesario diferenciar la dependencia constitutiva de la dependencia causal. El internismo no tiene por qué negar la segunda. Es posible que para tener el concepto de “vaso” yo haya tenido que estar causalmente relacionado con algún vaso, por ejemplo que haya tenido que ver vasos. Ahora bien, eso no quiere decir que para saber el concepto que yo tengo de vaso haya que mirar fuera de mi mente, inspeccionar la naturaleza de los vasos o el uso en el lenguaje de la palabra vaso no aportará ninguna información sobre el concreto que yo tengo de vaso. Eso es lo que defiende precisamente el internismo intencional.

17 Mirar (Moya, 2004, cap. 8, p. 155)

si no que parece reforzarlo.

7c) 7 y 8 se ve respaldado por la siguiente intuición: un sujeto y una réplica exacta de él tendrían los mismos conceptos y por tanto los mismo contenidos mentales

Un hecho que refuerza 8 es otra intuición que suele encontrarse tras la naturaleza de los estados intencionales y es que éstos determinan causalmente nuestro comportamiento¹⁸. Es de suponer que una persona y su réplica exacta se comportarían del mismo modo por lo que, si queremos explicar psicológicamente su comportamiento, hemos de individualizar los contenidos mentales desde una postura internista, hemos de analizar su contenido mental estrecho. Llamaremos “contenido mental estrecho” al contenido mental que resulta de una estrategia de individualización internista. Es decir, al contenido mental cuya naturaleza es explicada mediante factores estrictamente internos al individuo. En contraposición llamaremos “contenido mental amplio” a el contenido mental que resulta de una estrategia de individualización que apela no sólo a condiciones internas al individuo si no también externas (mundo, instituciones sociales..)

En conclusión, bajo la etiqueta “internismo intencional” englobaremos en el presente trabajo a todas las teorías y posturas acerca de lo mental que subscriben 8 (y por tanto 7). Es decir, a todas aquellas posturas que subscriben que la naturaleza última de todos nuestros contenidos intencionales depende solo de factores internos al sujeto (ya sea a la mente del sujeto o, como ocurre con las posturas materialistas, al cerebro del sujeto). Es decir, según el internismo intencional las condiciones necesarias y suficientes que hacen que nuestros contenidos representacionales sean los contenidos que son y no otros, es decir, las condiciones necesarias y suficientes que determinan el tipo de contenido intencional son siempre condiciones que no dependen de factores externos al sujeto pensante.

Como hemos dicho, las etiquetas “internismo y “externismo” intencional, surgen de la crítica que han hecho ciertas teorías (las teorías externistas) a ciertos presupuestos de la teoría tradicional de la mente. Llamaremos “externismo intencional” a las teorías de la

18 Por falta de tiempo y espacio, no consideraremos en este trabajo el fascinante debate en torno a la causalidad de lo mental. Para una exposición introductoria del tema remito a (Moya, 2004, tercera parte)

mente que recientemente han criticado las tesis 7 y 8 de la teoría tradicional. Es decir, de acuerdo con el externismo intencional, los factores internos al individuo no son suficientes para caracterizar el contenido intencional de todos nuestros pensamientos, es necesario, por tanto, tener en cuenta factores externos al sujeto. Finalizaremos este apartado con unas últimas consideraciones:

En la caracterización que hemos hecho de la teoría tradicional de la mente, hemos hecho especial hincapié en las propiedades semánticas de los contenidos intencionales. La razón de esto es obvia. Lo que define, como indicamos en el primer apartado, los contenidos intencionales es su capacidad de representar estados de cosas. Las propiedades representacionales (semánticas) de la mente permiten explicar, como hemos dicho, nuestro modo de pensar, de hacer inferencias, razonar.. pero no sólo eso. También parecen, intuitivamente, un factor clave a la hora de explicar nuestro comportamiento: como dice Burge¹⁹ actuamos de acuerdo a lo que creemos, deseamos...es decir, actuamos de acuerdo al objeto o contenido intencional de nuestros estados intencionales, de acuerdo a como nos representamos el mundo²⁰.

Por otro lado, si esto es así y las propiedades semánticas son constitutivas de los contenidos intencionales (tesis 2). el internismo intencional se ve entonces ante el reto de hacer esto compatible con las tesis 7 y 8. Es entonces necesario para la teoría internista que las propiedades semánticas de los contenidos intencionales (sean estos conceptos simples o contenidos más complejos) sean propiedades derivadas de propiedades internas a la mente del sujeto, o se explique en términos de éstas.

Por otro lado, nótese que cuando hemos dicho que las propiedades semánticas son constitutivas de los estados intencionales, tal como hemos utilizado el término “constitutiiva” a lo largo de la exposición de las tesis de la teoría tradicional de la mente, nos hemos estado refiriendo a condiciones necesarias (no suficientes). La razón de esto es simple: para que un contenido intencional sea un contenido intencional ha de

19 Por ejemplo en p. 361 . “Mind-Body Causation and Explanatory Practice,” in Heil and Mele (eds.) *Mental Causation*, Oxford: Clarendon Press, pp. 97–120. en (Burge 2007)

20 En la medida en que aceptamos que actuamos de acuerdo a motivos o razones, es obvio que tales motivos o razones tienen propiedades representacionales y hemos, por tanto, de explicar el comportamiento de alguien en base a dichas razones. Por ejemplo, si tengo sed y creo que hay agua en la nevera, el contenido representacional de mi creencia (el que yo me represente el agua en la nevera) es una razón por la que abro la nevera para coger agua.

tener propiedades representacionales, las propiedades representacionales son constitutivas del contenido intencional (hemos dicho que esto es lo que los caracteriza), sin embargo, por el hecho de que algo tenga propiedades representacionales, no quiere decir que ese algo sea un contenido intencional. Un cuadro o el lenguaje pueden tener propiedades representacionales y no son por ello contenidos intencionales.

Si bien a la hora de hablar de las propiedades semánticas lo hemos hecho en términos de propiedades constitutivas en el sentido de necesarias, esto no ha sido así en las tesis 7 y 8. En efecto, lo que caracteriza a las teorías internistas es que todas ellas consideran que los contenidos intencionales se reducen a propiedades internas del sujeto (de su mente o cerebro), Dichas propiedades no son sólo necesarias sino suficientes para caracterizar todos los tipos de contenidos intencionales (los contenidos intencionales se deben individualizar completamente mirando dentro del sujeto) así, las teorías internistas consideran que todos los contenido intencionales tienen un “contenido estrecho”. Por el contrario las teorías externistas tienen en común (y quizá sólo tienen en común) la negación de 8 (y por tanto de 7 si tenemos en cuenta 4), es decir, definiremos las teorías externistas como aquellas que afirman que no todos los contenidos intencionales son estrechos; es decir, (al menos) algunos contenidos intencionales son amplios. Cuántos contenidos intencionales no son estrechos, es decir, cuantos tengan un contenido amplio y no serán reductibles a factores internos del individuo, dependerá de la teoría externista que estemos considerando (algunas teorías externistas han defendido que sólo algunos contenidos intencionales son amplios, otras han ampliado el alcance a todos los contenidos intencionales).

La afirmación externista que reza que hay contenidos intencionales cuya naturaleza no depende únicamente de factores internos al individuo (negación de 8), ha de entenderse como que la naturaleza de un contenido intencional, las condiciones necesarias y suficientes que hacen que un contenido intencional sea del tipo que es no se encuentran únicamente dentro del individuo. Es decir, algunas de estos factores que determinan (individualizan) el contenido intencional son externos y dependen de factores contextuales tales como el lenguaje, el mundo, las instituciones, la cultura (qué tiene más peso dependerá de la teoría externista que se suscriba), lo cual, no excluye que también participen factores internos al individuo. Es decir, no quiere decir que para el externismo necesariamente todos los por factores que caracterizan la naturaleza de un

tipo de contenido mental sean externos, sino más bien que al menos alguno de ellos es externo.

Nótese que la caracterización que hemos hecho deja abierta la posibilidad de denominar con la etiqueta de “externistas” a teorías muy diversas y de muy distinto alcance. Así, parece que es externista desde una postura que considere, como hemos indicado, que sólo algún tipo de contenido es amplio, hasta una postura que afirme que todos los contenidos mentales son amplios. Incluso parecen externistas aquellas posturas que admitan una doble naturaleza del contenido²¹. En lo que sigue haremos una clasificación de las posturas externistas en base a dos criterios: En función de su alcance y en función de su fuerza²².

En función de su alcance: Diremos que una postura externista es de corto alcance si sólo admite la naturaleza amplia del contenido para unos pocos tipos de contenidos intencionales, y, diremos que una postura externista es de largo alcance si admite que la mayor parte o todos nuestros contenidos intencionales tienen una naturaleza amplia.

En función de la fuerza: Diremos que una postura externista es débil, si, para los tipos de estados mentales a los que se aplica, admite no sólo una estrategia de individualización externista si no también una internista (en el sentido en que podemos dar cuenta de su contenido estrecho), es decir, admite una doble naturaleza del contenido para los casos en que se aplica. Por el contrario diremos que una postura externista es fuerte si, para los casos en que se aplica, sólo admite una individualización externista del contenido, es decir, tales casos tienen un contenido amplio pero no un contenido estrecho.²³

En este trabajo exploraremos algunos de los argumentos más influyentes que se han dado desde la tradición analítica en favor de la estrategia de individualización externista de los contenidos intencionales. Es necesario tener en cuenta que dichos argumentos tienen sus antecedentes, no en el terreno de la discusión sobre los contenidos mentales, sino en el debate sobre la referencia que se estaba produciendo en el campo de la filosofía del

21 Explicaremos esto más adelante

22 Esta clasificación está parcialmente inspirada en la sección 2.3 de (Brown, 2011)

23 Esta clasificación se verá más abajo con más detalle así que pido paciencia al lector.

lenguaje.

En efecto, los argumentos de Putnam y Kripke basados en experimentos mentales que utilizan con objeto de defender las llamadas “teorías de la referencia directa”, han sido considerados como argumentos en favor de las tesis externistas del significado (sobre todo en el caso de Putnam). Algunos autores han llegado incluso a atribuir los experimentos mentales de Putnam una fuerza tal que servirían como argumentos en favor no sólo del externismo semántico, sino también del externismo intencional²⁴

Dada, por tanto, la influencia que estos argumentos han tenido sobre posteriores argumentos en favor del contenido amplio de los contenidos representacionales, un trabajo que pretenda estudiar los segundos deberá acercarse primero a la semilla que suponen los argumentos de Putnam y Kripke. Por este motivo, a continuación pasaremos a analizar el argumento de Putnam tratando de averiguar en qué medida dicho argumento supone un ataque al internismo semántico y plantea el externismo como alternativa. Tras esto, analizaremos los argumentos de Burge en favor del externismo en el terreno de lo mental y trataremos de identificar los elementos comunes y dispares a ambos argumentos con el objetivo último de entender cuál es la función y la fuerza que debemos atribuir a los mismos.

3. Antecedentes del externismo intencional: el externismo semántico²⁵

3.1 Teoría tradicional del significado

En el artículo *El significado de “significado”*²⁶ Hilary Putnam se propone criticar lo que él llama “la doctrina tradicional del significado”, para a continuación, esbozar una alternativa.

Se ha considerado que la crítica de Putnam es una continuación de las críticas de Kripke a las teorías descriptivas de la referencia. Sin embargo, esto no es del todo correcto. Tal como Putnam caracteriza la “doctrina tradicional”, esta puede considerarse el paradigma

24 Ver “The classic arguments for externalism” en (Lau y Deutsch, 2014)

25 Para el desarrollo de este apartado me he basado principalmente en (Putnam, 1975)

26 (Putnam, 1975)

o una idealización del internismo semántico y no de cualquier teoría descriptiva de la referencia²⁷. En este apartado explicaremos en que consiste dicha teoría del significado. Seguiremos, para ello, un proceso análogo a lo que hemos hecho en el apartado anterior con la caracterización idealizada de la teoría tradicional de la mente.

Putnam afirma que, de acuerdo con la “doctrina tradicional del significado”, podemos hablar de dos aspectos del significado, o dos sentidos de la palabra “significado”: la intensión y la extensión.

La extensión de un término se define como el conjunto de cosas de las que podemos predicar el término de un modo verdadero. Así, el término “conejo” se aplica o predica de y sólo a/de los conejos por lo que la extensión de “conejo” es el conjunto de todas las cosas de las que podemos decir que son conejos.²⁸

Sin embargo, como Frege indica en “sentido y referencia”²⁹, hay términos cuya extensión es la misma, pero que sin embargo difieren en cuanto a su significado. (por ejemplo: “criatura con riñones” y “criatura con corazón”. Esto es así, como indica Frege, porque difieren en el otro aspecto del significado, la intensión (o sentido)

La intensión, tal como la define Putnam, se identifica en la doctrina tradicional con el concepto o contenido intencional en sentido estrecho³⁰; es decir, tal como lo hemos definido en 8 (Esto permite explicar la diferencia entre “criaturas con corazón” y “criaturas con pulmones”: tienen asociados contenidos intencionales distintos (recuérdese anteriormente lo dicho acerca de esto)).

27 La razón por la que creo esto, es que Putnam dice dirigir su crítica a aquellas posturas que asumen el llamado solipsismo metodológico a la hora de caracterizar los estados psicológicos y asumen que conocer el significado es estar en un determinado estado psicológico de este tipo. No está claro que aquí caigan todas las teorías descriptivas de la referencia. Si bien, Putnam incluye en su crítica a Frege (Putnam, 1975, p. 352), Burge ha argumentado varias veces que Frege es un claro externista (o anti-individualista) por lo que no cae en la crítica de los argumentos basados en experimentos mentales. Ver por ejemplo, (Burge, 2007, p. 5)

28 Putnam afirma que ya esta noción de la teoría tradicional es problemática por sí misma. Resulta muchas veces complicado demarcar claramente los límites de la extensión de un término. Es decir, es dudoso en muchos casos qué elementos pertenecen al conjunto y cuáles no. (Putnam, 1975, p. 347)

29 (Frege, 1952)

30 En efecto, Putnam afirma que la noción de concepto o contenido mental que se asume desde la doctrina tradicional, se identifica con la noción de estado psicológico que surge de adoptar el presupuesto del solipsismo metodológico. Ver (Putnam, 1975, pp. 350-353)

Así mismo, de acuerdo con la doctrina tradicional, la intensión de un término es más básica que la extensión de tal modo la primera determina a la segunda pero no a la inversa, la extensión de un término no determina su intensión. Es decir, dos conceptos pueden tener la misma extensión y diferir en su intensión, sin embargo, dos términos no pueden tener la misma intensión y diferir en la extensión.

La última característica que Putnam asocia a la teoría tradicional del significado y que tiene relación con lo que acabamos de decir, es que , según dicha teoría, conocer el significado de un término es estar en un determinado estado psicológico. Podemos, en conclusión, condensar todo o que hemos dicho hasta ahora de la teoría tradicional del significado (internismo semántico) en las siguientes tesis:

Tesis 9. El significado tiene dos aspectos, intensión y extensión

Tesis 9.1 La intensión de un término es el contenido intencional tal como se ha descrito arriba asociado a ese término

Tesis 9.2 La extensión de un término es el conjunto de cosas de las que es correcto aplicar el término o que satisfacen predicar ese término de ellas

Tesis 10. La intensión de un término determina su extensión (pero no a la inversa)

Tesis 11. En el lenguaje, la intensión de un término se podría explicitar con una definición analítica. Una definición analítica, es una descripción que formula las condiciones necesarias y suficientes para la correcta aplicación del término³¹

Es fácil comprobar que lo que hace la teoría tradicional del significado que expone Putnam, en último término no es otra cosa que reducir el lenguaje (el significado) a la mente, a los contenidos intencionales (entendidos tal como los hemos caracterizado en la visión idealizada del anterior apartado) y explicar todo tal como se entiende desde la teoría clásica de los contenidos intencionales. Así, una vez reducido el significado a los contenidos intencionales, las tesis que acabamos de exponer no son más que

31 Ver apartado “1 definitional structure” en (Lawrence, Margolis, 2002)

reiteraciones de propiedades que antes atribuimos a los contenidos intencionales, pero vistas desde la óptica del lenguaje. Por ejemplo, las tesis 9 y 10 no son más que una reiteración de las tesis 5 y 6. El internismo semántico se caracteriza por reducir el significado a los contenidos intencionales entendidos en sentido estrecho (tal como hemos dicho que los entiende el internismo intencional) y explica, por tanto, las propiedades del significado, por ejemplo las propiedades representacionales, en términos de las propiedades representacionales de los contenidos intencionales (en sentido estrecho). De 9 y 10 se sigue por tanto la tesis general de que, para el internismo semántico, todos los aspectos del significado e un término vienen determinados por los contenidos mentales (en sentido estrecho):

Tesis 12: El significado de un término viene completamente determinado por los contenidos intencionales en sentido estrecho, es decir tal como son caracterizados según el internismo intencional (esto se sigue de 9, 9.1 y 10)

De lo que acabamos de decir, se desprende que la teoría internista del significado ha de aceptar necesariamente una teoría internista de los contenidos intencionales. Ese es el motivo por el que hemos numerado las tesis del internismo semántico como una continuación de las tesis del internismo intencional. Sin embargo, esto no debe llevarnos a error ya que la situación no se reproduce a la inversa. Es decir, tal como hemos caracterizado el internismo semántico, este presupone los contenidos intencionales en sentido estrecho, pero el internismo intencional, no presupone ninguna teoría del significado (en las tesis del 1 al 8 no hemos hecho referencia al significado en ningún momento). Esto quiere decir que el internismo intencional no va necesariamente ligado a una tesis internista del significado. Del primero no se desprende el segundo (aunque sí a la inversa).

Bien es cierto que la mayor parte de los autores que han defendido tesis internistas de los contenidos mentales, también han defendido un internismo semántico³², pero esto no es más que una contingencia histórica. Alguien, de un modo perfectamente coherente, puede subscribir una tesis internista de la mente y una tesis externista del significado. Para ésto sólo haría falta romper la relación entre lenguaje y mente (por ejemplo

32 El propio Putnam pone de ejemplo a autores como Carnap o Russell (Putnam, 1975 p. 352)

negando 9.1.). Nuestro filósofo podría pensar que las condiciones necesarias y suficientes que hacen que un contenido mental sea el contenido mental que es, son internas al sujeto y al mismo tiempo afirmar que el lenguaje no tiene nada que ver con la mente.

Volviendo a la caracterización que hemos hecho del internismo semántico, Putnam afirma que hay un corolario de dicha teoría, y es el siguiente:

Corolario 13. Conocer el significado de un término (en el sentido de intensión) no es otra cosa que estar en un determinado estado psicológico.

Para Putnam, 13 no es más que una reformulación de la tesis 9.1. Es decir, si el significado son los contenidos intencionales, entonces conocer el significado de un término implica estar en el contenido intencional que se asocia a ese término. Esto se ve reforzado por el postulado de la teoría tradicional de la mente, que Putnam parece asumir, según el cual tenemos conocimiento directo y completo de nuestros estados intencionales³³ (7a).

3.2 Crítica de Putnam a la teoría tradicional

Hemos definido el internismo semántico, como el conjunto de teorías que explican todo el significado a través de los contenidos intencionales en sentido estrecho (la conclusión 12). Las características que antes hemos expuesto no son más que el desarrollo de un modelo idealizado del internismo semántico. Supongamos ahora que queremos negar el internismo semántico, ¿Cómo podemos hacer tal cosa? Básicamente tenemos dos opciones: la primera de ellas sería la que ya hemos apuntado antes, establecer una separación entre lenguaje y mente. Es decir, seguir manteniendo que los contenidos intencionales son estrechos, pero negar que estos den cuenta del significado. La otra opción es mantener la relación entre lenguaje y mente, mantener que los contenidos intencionales determinan el significado pero negando que las condiciones de

33 Esto que acabamos de decir es una simplificación. Putnam aclara que, estrictamente, conocer el significado de un término no es meramente estar en el estado intencional que se asocia a ese término, si no saber, además que esa intensión está asociada a ese término. “Porque conocer el significado de A no es tan solo “captar la intensión” de A, con independencia de que pueda significar esto; es también saber que la “intensión” que uno ha “captado” es la intensión de A”. Sin embargo, para nuestros propósitos podemos simplificar y omitir esta aclaración ya que no marcará ninguna diferencia. (Putnam, 1975 p. 352)

individualización de los estados representacionales sean exclusivamente internas a la mente o cerebro del individuo (tesis 8), es decir, mantener un externismo intencional. ¿Cuál es la estrategia de Putnam? Por el momento dejaremos esta pregunta en un segundo plano y la retomaremos tras exponer su crítica al internismo semántico.

De un modo general, la crítica de Putnam a la teoría tradicional procede como sigue. Putnam afirmará que hay dos tesis centrales a la teoría tradicional del significado que son incompatibles ya que tienen consecuencias inaceptables. Para mostrar esto, Putnam desarrolla varios experimentos mentales que servirían de contraejemplo a las consecuencias que se siguen de dichos presupuestos. En este apartado trataremos de reconstruir el argumento de Putnam con un poco más de detalle.

Como hemos dicho, Putnam afirma que la teoría tradicional del significado descansa sobre dos presupuestos no cuestionados, cito directamente:

(I) Que conocer el significado de un término no es sino cosa de estar en un cierto estado psicológico (...)

(II) Que el significado de un término (en el sentido de “intensión”) determina su extensión (en el sentido de que mismidad de intensión implica mismidad de extensión)³⁴

El primer presupuesto se correspondería con la tesis 13 que hemos mencionado (que era a su vez un corolario de 9.1). El segundo supuesto se corresponde con la tesis 10. De la aceptación de ambos supuestos se seguiría según Putnam la siguiente consecuencia: dado que la intensión de un término se identifica con un estado psicológico (por el supuesto 13), y dado que la intensión determina la extensión (10), si dos personas se encuentran en el mismo estado psicológico con respecto a un término, necesariamente ese término tendrá la misma extensión en sus idiolectos. O lo que es lo mismo, no es posible que dos personas se encuentren en el mismo estado psicológico con respecto a un término y dicho término difiera en extensión.

Los experimentos mentales que Putnam desarrolla en *El significado de “significado”* se

34 (Putnam, 1975 p. 350)

orientan precisamente a negar esta consecuencia, lo que demostraría que al menos uno de los dos supuestos ha de ser incorrecto, y por tanto la teoría tradicional del significado (si entendemos una teoría como un conjunto de tesis) también sería incorrecta. Una posible reconstrucción del argumento genérico sería así:

(1) Si entender el significado de un término (la intensión de un término) es estar en un determinado estado psicológico y la intensión de un término determina su extensión, entonces no es posible que dos personas se encuentren en el mismo estado psicológico con respecto a un término y dicho término difiera en su extensión.

(2) Es posible que dos personas se encuentren en el mismo estado psicológico con respecto a un término y la extensión del término difiera en sus respectivos idiolectos

(3) No es correcto afirmar que entender el significado de un término sea estar en un determinado estado psicológico y que la intensión de un término determine la extensión de un término (al menos una de las dos cosas a de ser falsa)

(4) Si la teoría tradicional del significado es correcta, entender el significado de un término es estar en un determinado estado psicológico y la intensión de un término determina su extensión,

(5) La teoría tradicional del significado es incorrecta

La estructura general del argumento de Putnam es un *Modus Tollens*. La primera premisa, está justificada arriba (desarrolla una consecuencia de la teoría internista del significado), y la tercera premisa, no es una premisa sino la conclusión del argumento. La fuerza del argumento dependerá por tanto del grado en que Putnam sea capaz de justificar la segunda premisa. Es decir, Putnam ha de mostrar que es posible que dos personas se encuentren en el mismo estado intencional con respecto a un término y que la extensión de dicho término varíe en sus respectivos idiolectos.

3.3 Experimentos mentales de las tierras gemelas

Para mostrar la premisa (2), Putnam hace uso de varios experimentos mentales. A

continuación expondremos algunos de ellos para posteriormente proceder al análisis de las conclusiones que de ellos podemos extraer³⁵.

a) Supongamos que en algún lugar perdido de la galaxia hay un planeta al que llamaremos Tierra Gemela. La Tierra Gemela es prácticamente una copia exacta de la Tierra. Hasta tal punto es así que podemos suponer que hay una copia exacta de todas las personas que viven en la Tierra, de hecho también hay habitantes en este planeta que hablan español, sin embargo, el español que se habla en la Tierra Gemela tiene alguna peculiaridades que lo distinguen del español que se habla en la Tierra como veremos a continuación.

Una de las peculiaridades que distinguen a la Tierra Gemela de la Tierra, es que en la Tierra Gemela el líquido que los hablantes llaman “agua” no es H₂O sino uno muy diferente cuya fórmula química es muy larga y complicada, abreviaremos dicha fórmula con la nomenclatura XYZ. Dicho líquido, salvo en su composición química, es indistinguible del agua. Sabe como el agua, quita la sed, fluye por los ríos y océanos, cae del cielo cuando llueve..

b) Putnam afirma que si una nave espacial de la Tierra tripulada por terrestres aterrizara algún día sobre la Tierra Gemela, los terrestres al principio supondrían que el término “agua” tiene el mismo significado en la Tierra Gemela que en la Tierra. Sin embargo, tras investigar su composición, rectificarían y enviarían el siguiente informe a la Tierra:

“En la Otra Tierra, la palabra “agua” significa XYZ”

Lo mismo ocurriría en la situación inversa. Es decir, si una nave de la Tierra Gemela aterriza sobre la Tierra. Al cabo de un tiempo los tripulantes mandarían un informe: “En la Tierra la palabra “agua” significa H₂O.

Es claro, dice Putnam, que la palabra “agua” tiene distintos significados en las dos tierras ya que, si un aspecto del significado es la extensión (tesis 9 de la teoría clásica del significado), es claro que esta varía en el español de la Tierra y el español de la

35 Los experimentos que resumimos a continuación se pueden encontrar en (Putnam, 1975 pp. 353-537)

Tierra Gemela. En uno es el conjunto de la totalidades formada por moléculas de H₂O, y en otro es el conjunto de las totalidades formadas por moléculas de XYZ.

b) Putnam advierte lo siguiente. Imaginemos que Oscar es un tipo terráqueo que habla español y tiene un duplicado exacto de él en la tierra gemela (llamémosle Oscar Gemelo). Oscar y Oscar Gemelo, son idénticos a nivel físico, neuronal, han tenido una vida perfectamente paralela, por lo que tienen las mismas creencias, deseos, qualias, aspiraciones, sensaciones, etc. Salvo por el lugar en que habitan, Oscar y Oscar Gemelo son idénticos y han tenido vidas idénticas en todos los aspectos. Asumamos además que tanto Oscar como Oscar Gemelo no conocen la estructura de eso que ellos denominan “agua”. Putnam afirma que Oscar y Oscar Gemelo se encontrarían, en dicha situación hipotética, en el mismo estado psicológico y, sin embargo entenderían el significado del término agua de modo distinto ya que al hablar Oscar de agua, éste se estaría refiriendo a H₂O, mientras Oscar Gemelo se referiría a XYZ. De nuevo, dado que la extensión de sus términos es distinta, el significado del término “agua” es distinto dependiendo de quién de los dos lo pronuncie.

c) Putnam lleva todavía más lejos el experimento mental y nos pide que que retrocedemos a 1750 (época en la que todavía se desconoce la composición del líquido llamado “agua” tanto en la Tierra como en la Tierra Gemela). Putnam afirma que cuando un hablante de la Tierra, y su réplica en la Tierra Gemela hablen de “agua”, ambos significarán cosas distintas a pesar de encontrarse en estados psicológicos idénticos y a pesar de que sus respectivas comunidades científicas no conocen aún la fórmula química del elemento que llaman “agua”.

d) Putnam construye varios experimentos mentales en la misma línea de los anteriores, uno de los más comentados ha sido el que se refiere a los términos “olmo” y “haya”. Putnam advierte (sin imaginar tierras gemelas) que es posible que, debido a su ignorancia en botánica, alguien se encuentre en el mismo estado psicológico al utilizar la palabra “olmo” y la palabra “haya”. Sin embargo, la extensión de los dos términos en el idelecto de dicha persona es distinta. La palabra “olmo” tiene como extensión el conjunto de todos los olmos y la palabra “haya” tiene por extensión el conjunto de todas las hayas. Esto muestra que, de nuevo a mismo estado psicológico puede no corresponderle misma extensión, y por lo tanto teniendo el mismo concepto en la cabeza,

puede significar cosas distintas.

Estos ejemplos, y otros que desarrolla en el texto, vendrían a demostrar lo mismo: Es posible que dos personas se encuentren en el mismo estado psicológico, tengan en mente los mismos conceptos o contenidos intencionales, y a pesar de todo la extensión, y por tanto el significado de los términos, varíe. Estos ejemplos demostrarían la premisa (2) del argumento expuesto arriba contra las consecuencias que se siguen de dos de los presupuestos básicos en que se basa la teoría internista del significado. Por lo tanto, la teoría internista es falsa. Los significados no están determinados por los contenidos intencionales (estrechos). En palabras de Putnam:

¡Córtesé el pastel como se quiera, pero los “significados” no pueden estar en la cabeza!³⁶

3.4 Teoría alternativa de Putnam

Una vez demostrada la inviabilidad de la teoría tradicional del significado, Putnam propone una teoría alternativa que pueda dar cuenta correctamente de los casos antes explicados. A continuación revisaremos muy brevemente dicha teoría.

¿Cómo son posibles los casos de los experimentos mentales? ¿Cómo es posible que dos personas se encuentren en el mismo estado psicológico pero, la extensión de sus términos y por tanto el significado varíe? Para explicar este hecho, Putnam propone lo que él llama la “hipótesis de a división del trabajo lingüístico”.³⁷

En efecto, Putnam afirma que un persistente error de la teoría tradicional, fue ignorar el aspecto social y cooperativo del lenguaje. Para explicar su propuesta, Putnam pide que pensemos en muchas de las actividades cooperativas que realizamos los seres humanos. Si pensamos, por ejemplo, en el mundo alrededor del oro, vemos que hay gente que lo vende, gente que lo lleva puesto en forma de anillos, gente que se dedica a distinguir si algo es o no oro... No es necesario que todo el mundo que lleve oro, sea capaz de distinguir si algo es o no oro, igual que tampoco es necesario que el experto en distinguir si algo es o no oro, sepa del precio del oro. Si una persona quiere saber si su

36 (Putnam, 1975 , p. 357)

37 (Putnam, 1975 pp. 357-359)

anillo es o no de oro, basta con que consulte al experto. Algo similar, dice Putnam, ocurre con el significado. No es necesario que todas las personas que hacen uso de la palabra “oro” sepan distinguir el oro, delegan ese trabajo a una comunidad de expertos. Es decir, en muchos de nuestros conceptos, delegamos en una comunidad de hablantes (que no sería otra cosa que un subconjunto de la comunidad total de hablantes) para que estos determinen la referencia de nuestros términos. Igual que gran parte del trabajo mundano, el lenguaje tiene un aspecto cooperativo y para que una comunidad signifique algo con un término, no es necesario que todos los rasgos de su significado sean completamente conocidos por cada uno de los individuos de la comunidad. Basta con que dichos rasgos pertenezcan al significado social del término

En toda comunidad lingüística se da la clase de división del trabajo lingüístico acabada de describir: es decir, toda comunidad lingüística posee al menos algunos términos cuyos correspondientes “criterios” sólo los conoce un subconjunto de los hablantes que los aprendieron y cuyo uso por parte de los restantes depende de una cooperación estructurada entre éstos y los hablantes de los subconjuntos relevantes.³⁸

Esta hipótesis explica el fracaso, mostrado en los experimentos mentales, de los supuestos 10 y 13 de la teoría tradicional, de acuerdo con los cuales no era posible que dos personas compartieran el mismo estado psicológico y refirieran a cosas distintas.

Podemos resumir esta discusión señalando que hay dos clases de herramientas en el mundo: herramientas, como un martillo o un destornillador, que pueden ser usadas por una sola persona; y herramientas como un barco a vapor, que, para ser usadas, requieren la actividad cooperativa de un cierto número de personas. De las palabras se ha pensado en demasía que responden al primer modelo de herramienta.³⁹

El mecanismo de la división del trabajo lingüístico que acabamos de describir, es posible gracias a ciertos rasgos del significado que Putnam desarrolla en su teoría⁴⁰ y

38 (Putnam, 1975, p. 359)

39 (Putnam, 1975, p. 360)

40 El núcleo del desarrollo puede encontrarse en las (Putnam, 1975, pp.360-366)

que enlazan directamente con la naturaleza del significado de los términos naturales
¿Cuál es esta naturaleza?

Volvamos por un momento al ejemplo de la tierra gemela. Lo que las intuiciones de dichos ejemplos mostraban, es que era posible que dos personas se encontraran en el mismo estado psicológico, y sin embargo, se refirieran con el lenguaje a cosas distintas. Es decir, las intuiciones respaldan la premisa (2) del argumento del que partíamos. Por tanto, podemos concluir (3). Si las intuiciones de los experimentos mentales son correctas, entonces como dice Putnam:

La teoría de que (1) las palabras tienen “intensiones”, que son algo parecido a los conceptos vinculados a las palabras de los hablantes; y que (2) la intensión determina la extensión, no puede ser verdadera en lo que toca a las palabras que designan clases naturales como “agua”⁴¹

Parece, por tanto, que si queremos desarrollar una teoría del significado que sea minimamente compatible con las intuiciones de los experimentos mentales, y queremos salvar alguno de los presupuestos tradicionales, tenemos al menos dos posibilidades. O negamos el primer presupuesto, o negamos el segundo⁴². Es decir, de los ejemplos podemos concluir que al menos uno de los dos presupuestos no es aceptable.

Si optamos por negar la tesis 10 de la teoría tradicional habríamos de afirmar que términos como “agua” tiene un significado “relativo-constante”. Es decir, Oscar, y Oscar Gemelo significarían los mismo con la palabra agua (en el sentido en que tienen la misma intensión) pero la extensión de sus términos sería relativa al mundo en que se encontraran y no estaría determinada por propiedades intrínsecas a el sujeto. Es decir, seguiríamos considerando que la intensión (entendida como equivalente al contenido intencional estrecho) es parte del significado pero ésta no determina la extensión de los términos.

41 (Putnam, 1975, p. 365)

42 Podemos pensar en una tercera posibilidad que consistiera en una teoría del significado que negara ambos presupuestos. Putnam no considera dicha posibilidad (al menos en El significado de “significadp”). Sin embargo, podemos imaginar que una teoría de este tipo supondría cambiar radicalmente la teoría tradicional el significado y proponer algo completamente nuevo.

La segunda opción, consiste en rechazar en su lugar la tesis de que la intensión se identifica con el contenido mental del sujeto, es decir, rechazar la tesis 9.1 de la teoría tradicional del significado. Esta vía como veremos, implica establecer una separación tajante entre lenguaje y mente.

En el significado de “significado” Putnam admite que ambas vías pueden ser compatibles con las intuiciones de los experimentos mentales⁴³ sin embargo, opta por la segunda y desarrolla una teoría del significado de los términos naturales que dé cuenta de ella. No nos detendremos a analizar a fondo dicha teoría ya que excedería los propósitos de este trabajo. Sin embargo podemos resumirla como sigue.

La idea general es que los significados de las palabras que designan clases naturales como oro, agua, u olmo, tienen un componente indéxico (aunque no totalmente indéxico como ocurre en el caso de palabras como “esto”, “eso” “yo”. Si fuera totalmente indéxico estaríamos en el caso de la primera posibilidad. La intensión se identificaría con un contenido mental pero ésta no determinaría la extensión.)

Imaginemos que tratamos de explicar a alguien qué es el agua mediante una definición ostensiva. Es decir, señalamos una muestra de agua y decimos “esto es agua”. ¿Cómo hemos de interpretar el significado de la palabra agua en este contexto?

Según Putnam, cuando hablamos de “agua” u otras palabras que designan clases naturales, el hablante supone que ese elemento ha de tener una estructura o naturaleza común (aunque la desconozca) y que todo objeto que comparta dicha naturaleza ha de ser agua. Así, cuando decimos la frase “esto es agua” el significado de agua ha de interpretarse como estar en una relación de mismidad (mismo elemento) con la entidad que señalamos: con “esto”. Es decir, agua es todo aquello que sea el mismo elemento, la misma clase (tenga la misma estructura última) que el agua de por aquí, el agua al que señalo (aquí está el componente indéxico).

Volvamos por un momento la mirada a los experimentos mentales: Si los analizamos, hay diferencias entre ellos. El primero de todos, no es realmente un caso problemático

43 Ver (Putnam, 1975, 353)

para la teoría tradicional del significado ya que cuando los tripulantes de la nave terráquea descubren que lo que llaman “agua” en la tierra gemela es XYZ, su estado psicológico es distinto al pensar en el “agua” de allí, al del “agua” de la Tierra. El segundo experimento y el cuarto, por el contrario, si suponen un problema para la teoría tradicional ya que Oscar y Oscar Gemelo se encuentran en el mismo estado psicológico pero (si aceptamos las intuiciones finales) el significado difiere. Lo mismo ocurre con el caso de “olmo” y “haya”. Estos dos experimentos respaldarían la premisa (2) del argumento. La explicación de Putnam de la “división del trabajo lingüístico” parece poder dar cuenta de ellos ya que aunque Oscar y Oscar Gemelo no difieran en su contenido mental estrecho, ya hay en sus sociedades respectivas diferencias respecto al significado social de la palabra “agua”. Sin embargo, la hipótesis de la división del trabajo lingüístico por sí sola (al menos tal como la hemos formulado), no permite dar cuenta del tercer experimento mental, según el cuál los hablantes de 1750 significan cosas distintas a pesar de que sus respectivas comunidades científicas desconozcan la estructura de lo que llaman “agua”.

Sin embargo, la teoría que acabamos de mencionar acerca del componente indécico de los términos que designan clases naturales, sí que permite ya dar cuenta de todos los casos. Incluido el tercero. En efecto, aunque las sociedades de la Tierra y la Tierra gemela lo desconozcan, el significado de la palabra “agua” es distinto solo que todavía no lo saben. Ambos elementos no están en una relación de “mismidad” por lo que cuando un hablante de la Tierra habla de “agua” en la extensión del término, a pesar de que nadie en su sociedad lo sepa, no entra la sustancia que los hablantes de la Tierra Gemela llaman “agua” ya que no es el mismo elemento que el agua de “por aquí” de la tierra (a la que designa el componente indécico de la palabra “agua”)

Esta teoría de los términos que designan clases naturales, en opinión de Putnam parece amoldarse bien a nuestras intuiciones del desarrollo científico⁴⁴, permite explicar que cada vez sepamos más de algo. Parece más intuitivo decir que hoy conocemos la estructura química del agua y que la gente de 1750 no, y por lo tanto hoy conocemos más del agua que lo que la gente de 1750 conocía sobre el agua, que decir que el significado de la palabra “agua” es distinto ahora al que el que tenían las personas de

44 Ver el apartado “seamos realistas” de (Putnam, 1975)

1750. Decimos que en el siglo X antes de Cristo no se conocía todavía la estructura química del oro, y no que significaban algo distinto de la palabra “oro”.

Esta teoría también permite explicar, dice Putnam, la falibilidad de la ciencia. Así podemos pensar que en el futuro se descubriera que la estructura química del agua no es H₂O. En ese caso diríamos que estábamos equivocados con respecto a la estructura química del agua (estábamos equivocados sobre la estructura química del elemento al que referíamos mediante el componente índice del significado) y no parece intuitivo decir que en ese caso diríamos que ha resultado que no había agua en el mundo.

Parece por tanto, que lo que podamos suponer en los ejemplos contrafácticos de mundos posibles, está en función de lo que de hecho ocurre en la tierra, en el lugar en el que está el elemento al que nos referimos mediante el componente índice del significado de los términos de clase natural.

La teoría de Putnam, se ha considerado, y el mismo la considera⁴⁵, una continuación de la teoría de Kripke acerca de los nombres propios⁴⁶. Estas teorías se encuadran en lo que se ha llamado “teorías histórico-causales de la referencia”, en el caso de Kripke, o teorías de la referencia directa. No entraremos aquí a explicar en detalle en qué consisten ni explicaremos tampoco la teoría de Kripke. Simplemente diremos que la principal diferencia entre las dos es el carácter más pragmático de la teoría de Putnam (parte del comportamiento lingüístico de los hablantes al designar clases naturales. Hacemos “como si” tuvieran una estructura interna, fueran un elemento del mundo) frente al más realista de Kripke (estas palabras hacen de designadores rígidos que nombran entidades metafísicamente reales del mundo).

En este apartado no hemos querido entrar al detalle de la teoría del significado de Putnam, sólo hemos querido hacer referencia a los aspectos de su teoría que dan cuenta de los casos de los experimentos mentales. Una vez hecho esto, estamos en situación de preguntarnos: ¿qué podemos concluir de los argumentos de Putnam y de su solución para el debate entre el internismo y el externismo semántico?

45 (Putnam, 1975, p. 362)

46 Para ver en detalle esta teoría remito a (Kripke, 1972)

En primer lugar, como hemos afirmado arriba, hay dos vías disponibles compatibles con los resultados de los experimentos mentales. 1) Seguir identificando el significado (en el sentido de intensión) con el contenido intencional estrecho y negar la tesis de que este determine la extensión del término.

Es fácil ver que esta vía implicaría hacer cambios radicales en la teoría tradicional de los contenidos intencionales que hemos caracterizado arriba y no sólo en la del significado (por ejemplo en el hecho de que las propiedades semánticas sean constitutivas de los contenidos o que los conceptos determinen las entidades que representan...) sin embargo no entraremos a analizar esta opción, aunque el coste de adoptarla parece (a simple vista) ser alto. Pero el hecho de que esta vía sea una opción, es relevante en la medida en que el internismo semántico e intencional pueden “salvarse” de los experimentos mentales. A costa, eso sí, de grandes modificaciones en la teoría tradicional del contenido mental.⁴⁷

Putnam prefiere, por el contrario, optar por la segunda vía. 2) Esta implica dejar de identificar cualquier aspecto del significado con los contenidos intencionales y conservar la tesis de la teoría tradicional del significado según la cual la extensión es un elemento constitutivo de este. Así, un cambio de extensión es, en esta vía, sin más, un cambio de significado⁴⁸ Lo que hace que las propiedades semánticas sigan siendo constitutivas del significado.

Por otro lado, antes afirmamos que, para poder mantener una tesis externista del significado, había dos vías. 1) Explicar el significado mediante los contenidos intencionales (como ocurre en la teoría tradicional del significado que hemos caracterizado) pero negando que haya que entender estos en sentido estrecho (negar las tesis 7 y 8 de la teoría tradicional de los contenidos intencionales). Es decir, explicar el

47 En este caso podríamos decir que los experimentos no mostrarían concluyentemente la incorrección del internismo semántico. Dado que desde esta postura, un cambio en la extensión no implica un cambio en el significado, se salvaría la intuición internista de que dos personas en el mismo estado psicológico han de significar lo mismo. Sin embargo, esto no debe en absoluto ser una motivación para que el internista comience los festejos. Aun si mantenemos esta vía, la teoría internista queda seriamente comprometida. ¿Hasta que punto es aceptable una teoría del significado en la que la extensión no sea un componente constitutivo del significado? ¿Hasta que punto es admisible una teoría que casi vacía al significado de sus propiedades semánticas? Una teoría de este tipo parece casi una contradicción en sus propios términos, una teoría del significado de este tipo no parece una teoría del significado.

48 (Putnam, 1975 , p. 365)

significado a través de lo mental, pero manteniendo una postura externista de la mente. O, 2) seguir manteniendo una teoría internista de la mente pero separar esta del significado. Lo que implicaría negar la tesis 9.1 de la teoría tradicional del significado y con ello, al menos las tesis 12 y 13. Esta segunda opción es la que parece mantener Putnam. Como dice en *el signifocado de "significado"*: los significados no están en la cabeza. El significado no tiene nada que ver con la mente del sujeto, aunque la mente del sujeto pueda caracterizarse mirando en la "cabeza" del sujeto.⁴⁹

Si traspasamos los criterios de clasificación que hemos hecho de las teorías externistas en el segundo apartado al significado, ¿Qué podemos concluir de la teoría de la teoría de Putnam? En primer lugar parece que el alcance es limitado. Se aplica a los términos de clase natural (puede que se aplique a más pero Putnam no se compromete con ello). Por otro lado, para los términos que se aplica, la teoría parece ser fuerte. Es decir, el significado de estos términos no puede entenderse en términos de contenidos mentales, (hemos dicho que el significado no tiene que ver con la mente según Putnam).

4. Burge y el anti-individualismo de los contenidos mentales

Hasta ahora hemos visto argumentos externistas significado y la postura concreta de Putnam, estos pasan por separar lenguaje y mente, Putnam sigue admitiendo contenidos intencionales estrechos (como hemos visto), sin embargo los argumentos de Putnam han servido para dar un paso más, e inspirar argumentos análogos en el terreno de la mente con el objetivo de argumentar en favor de una postura externista de los contenidos intencionales, es decir una postura externista de la mente y no sólo del lenguaje.

Burge es quizá el principal autor en hacer uso de esta estrategia para argumentar en favor de una postura externista de los contenidos intencionales, o como él la llama, "anti-individualismo", en contraposición al individualismo (o internismo acerca de los contenidos intencionales). En este sentido, el método que utiliza para argumentar en favor de dicha postura es completamente análogo al de Putnam; si bien, como veremos, las consecuencias que él extrae son radicalmente distintas.

Burge toma como punto de partida la tesis tradicionales de la mente (más o menos) tal

⁴⁹ Justificaremos y veremos a discutir esta posición más abajo.

como han sido descritas arriba ⁵⁰. Como dijimos, lo definitorio del internismo intencional (del individualismo, en términos de Burge) es que la naturaleza de los contenidos intencionales, las condiciones necesarias y suficientes que hacen que un contenido intencional sea el contenido intencional que es dependen únicamente de factores internos al sujeto (tesis 7 y 8). Por lo tanto, para individualizar los contenidos intencionales hemos de hacerlo en función de estas características.

Si esto es así, de aquí se sigue la consecuencia de que si dos personas son idénticas físicamente y en su historia mental (no intencional), entonces sus contenidos intencionales y por tanto sus estados intencionales (por la tesis 1) deben de ser idénticas (incluyendo sus propiedades semánticas). Los factores que no sean internos al sujeto, son irrelevantes para caracterizar los estados intencionales de un sujeto.

Los argumentos de Burge, tal como ocurría en el caso de Putnam, están orientados a aportar intuiciones que problematicen esta consecuencia. Así, vamos provisionalmente a presentar el argumento de Burge de modo análogo al argumento de Putnam; y, digo provisionalmente, porque resulta problemático hasta que punto podemos encuadrar el argumento de Burge (la totalidad de lo que pretende extraer de los experimentos mentales) meramente en un esquema como el siguiente. Discutiremos esto más abajo.

(1) Si aceptamos el internismo intencional, entonces hemos de aceptar que es imposible que dos personas sean idénticas y sus estados intencionales varíen.

(2) Es posible que dos personas sean idénticas y sus estados intencionales varíen

(3) El internismo intencional no es aceptable

Pasaremos entonces (de momento) a analizar los experimentos mentales y cómo estos (si es que lo hacen) justifican la segunda premisa del argumento igual que hicimos en el caso de Putnam.

4.1 Experimentos mentales

50 Ver (Burge, 1993)

El experimento mental más conocido⁵¹ de Burge es el que refiere al concepto de artritis. Burge lo divide en tres pasos:

Primer paso: Supongamos que una persona tiene un gran número de actitudes intencionales asociadas contenidos intencionales que contienen el concepto de “artritis”. Esta persona cree (correctamente) que ha tenido artritis durante años, cree que la artritis que padece en las muñecas y los dedos es más dolorosa que la que padece en los tobillos, cree que es mejor tener artritis que cáncer de de hígado, cree que la rigidez de las articulaciones es un síntoma de la artritis, etc. Sin embargo, cree falsamente que recientemente ha desarrollado artritis en el muslo.

Supongamos que dicha persona decide ir al médico para comunicarle su temor a que la artritis se haya extendido a su muslo. El doctor le contesta entonces que tal cosa no es posible ya que la artritis es específicamente una inflamación de las inflamaciones. Cualquier diccionario podría haberle dicho eso. El paciente sorprendido, le pregunta entonces qué es lo que puede tener en el muslo.

Segundo paso: Imaginemos una situación contrafáctica en la que el paciente ha tenido la misma vida, se ha comportado igual, su vida física y sus fenómenos mentales no intencionales (estados de conciencia, etc..) han sido exactamente los mismos. Igual que en el primer paso, el paciente desarrolla la predisposición a asentir ante la emisión “la artritis puede ocurrir en el muslo”. La única diferencia en esta situación contrafáctica concierne el entorno social del paciente: en la situación contrafáctica el término “artritis” tal como es usado en su comunidad, se aplica a varias enfermedades reumáticas, incluyendo la que la persona tiene en su muslo e incluyendo la artritis.

Tercer paso: El último paso es una interpretación del caso contrafáctico. Es razonable suponer (nuestra intuición nos dice que) en la situación contrafáctica el paciente carece de algunas (probablemente todas) las actitudes normalmente ligadas a cláusulas de contenido intencional que contengan el concepto de artritis. El paciente carece de la creencia de que tiene artritis en su muslo, que ha tenido artritis durante años, etc...

51 (Burge, 1979 pp. 104-106)

En el caso contrafáctico no podemos, dice Burge, adscribir al paciente ninguna cláusula de contenido intencional que contenga el término “artritis”. De modo análogo al caso del “agua” de la tierra gemela, resulta difícil imaginar cómo el paciente ha adquirido la noción de artritis si el término “artritis” en su comunidad contrafáctica no significa artritis⁵².

Our ascriptions of content clauses to the patient (and ascriptions within his community) would not constitute attributions of the same content clauses we actually attribute. For counterpart expressions in the content clauses that are actually and counterfactually ascribable are not even extensionally equivalent.⁵³

Este experimento mental y muchos otros que Burge comenta, mostrarían que es posible que los contenidos mentales de una persona (en este caso el paciente) difieran, mientras toda su historia física y mental (no intencional) considerada de manera aislada al resto del contexto social, permanece inalterada. El paciente del primer caso y del segundo caso son idénticos y, sin embargo, sus estados intencionales son distintos.

4.2 La interpretación de Burge de los experimentos mentales

¿Qué hemos de concluir de estos experimentos? La clave del experimento está es que, de acuerdo con las intuiciones que supuestamente se siguen de ellos, no atribuiríamos al paciente en el mundo actual y al paciente en el mundo contrafáctico el mismo contenido intencional. Diríamos, desde el análisis ordinario, que el paciente del mundo actual cree que tiene artritis en el muslo y que el paciente de la situación contrafáctica cree que tiene otra enfermedad, llamada “artritis” en su mundo contrafáctico, en el muslo.

Si aceptamos estas intuiciones, entonces la premisa 2 del argumento es válida y por tanto tenemos que concluir que el internismo intencional no es aceptable, y de ahí concluir que el externismo intencional sí lo es. De hecho, en los experimentos, teniendo en cuenta los elementos contextuales, sí que podemos individualizar los contenidos intencionales. De ahí que se sigan las intuiciones que acabamos de indicar.

52 En la situación contrafáctica estamos suponiendo también que no hay una palabra en el repertorio del vocabulario del sujeto que signifique artritis tal como se entiende en el mundo real.

53 (Burge, 1979 p. 106)

Ahora bien, aun si asumimos que los experimentos mentales son concluyentes para respaldar la segunda premisa del argumento y apoyar una postura no internista de la mente. El argumento inicial en tres pasos no dice nada acerca de la postura externista que debemos concluir. Es decir, ¿Qué tipo de externismo intencional podemos concluir de los experimentos? En el segundo apartado, vimos que las posturas externistas pueden variar mucho. Tanto en su alcance como en su fuerza. Si queremos analizar la fuerza de los experimentos mentales, hemos de analizar no sólo si sirven para negar o no el internismo intencional, si no ver que tipo de tesis externistas pueden o pretenden apoyar. Para ello hemos de preguntarnos por lo que podemos concluir de ellos y que dicen sus conclusiones sobre los dos criterios clasificatorios que mencionamos arriba. Es decir, hemos de preguntarnos dos cosas: i) ¿A qué tipo de contenidos pueden aplicarse sus intuiciones? Es decir ¿qué alcance tienen los argumentos? ¿A qué contenidos afecta? Y ii) para los contenidos que afecta ¿Qué hemos de concluir sobre esos contenidos?

Comenzaremos por la segunda pregunta: para los tipos de contenidos a que afecta, ¿Qué fuerza tienen los argumentos? Una opción sería interpretar sus intuiciones del siguiente modo: ordinariamente individualizamos los contenidos intencionales en base a una estrategia externista. Es decir, teniendo en cuenta aspectos contextuales (como el uso que hace la comunidad científica o los hablantes de un término) o externos a la mente del sujeto. Individuamos el contenido amplio de los sujetos. De ahí que nuestra intuición nos diga que en el caso del paciente real y el contrafáctico diríamos (ordinariamente) que tienen distintos contenidos.

Ahora bien, eso no excluye que de un modo “no-ordinario” podamos atribuirles un contenido estrecho común a ambos pacientes. ¿Qué tipo de contenido sería este? Quizás un contenido que podríamos parafrasear con una descripción (del tipo: inflamación de las articulaciones y el muslo) o por ejemplo, un contenido que sea algo así como una función del contexto a las condiciones de verdad similar a lo propuesto por Fodor⁵⁴, aunque eso implique modificar levemente alguna de las tesis de la teoría tradicional, mientras no se toquen las tesis 7 y 8. En cualquier caso, un contenido que nos permita mantener tesis clásicas de la explicación de la mente como el llamado funcionalismo o

54 Ver (Fodor, 1988. pp. 52-53)

la teoría de la superveniencia de las propiedades mentales sobre las físicas, por ejemplo.

Esta posición, supondría afirmar que los experimentos son compatibles con un externismo débil que de pie a admitir una doble naturaleza del contenido (amplio en el sentido ordinario y estrecho en un sentido no-ordinario) ¿Es esta la fuerza que Burge les atribuye? Creo que este no es el modo correcto de interpretar la posición de Burge. La fuerza que parece atribuir a los experimentos es mucho más fuerte que una postura como la que acabamos de describir. La razón es la siguiente. Allá donde pueden aplicarse los experimentos mentales, Burge parece asumir que sus intuiciones no sólo muestran la validez de aceptar una individualización externista del contenido, si no que ponen en cuestión y traen escepticismo sobre la propia estrategia de individualización que sigue el internismo, es decir sobre la propia plausibilidad de la noción de contenido estrecho para los casos en que los experimentos se aplican.

En la Post-Data de *Individualism and the Mental* Burge mantiene que concluyentemente los individuos en la situación actual y contrafáctica tienen pensamientos con contenidos intencionales (representacionales distintos) y que hay elementos externos que son constitutivos de los contenidos intencionales. Es necesario tener en cuenta esos elementos para individuar los contenidos intencionales⁵⁵. Precisamente lo que niega la tesis tradicional de la mente que hemos explicado en el segundo apartado.

En una línea similar, en “Other Bodies” afirma:

(...) But the arguments of “Individualism and the Mental” suggest that virtually no propositional attitudes can be explicated in individualistic terms. Since the intentional notions in terms of which propositional attitudes are described are irreducibly non-individualistic, no purely individualistic account of these notions can possibly be adequate⁵⁶

En el artículo “The two thought experiments reviewed” Burge es todavía más explícito en esta segunda interpretación de sus experimentos mentales:

55 (Burge, 1979 pp. 155-156)

56 (Burge, 1982, p. 99)

What do the thought experiments show? They show that intentional content of ordinary propositional attitudes, as indicated by obliquely occurring expressions in that-clauses, cannot be accounted for in terms of physical, phenomenal, causal-functional, computational, or syntactical states or processes that are specified nonintentionally and are defined purely on the individual in isolation from his physical and social environment. Intentional content, in the aforesaid sense, is not even supervenient on the nonintentional processes and states of an individual, insofar as these processes and states are “individualistically” described. Thus individualistic, functionalist, computationalist or physicalist accounts of ordinary intentional content fail in a systematic manner.⁵⁷

La postura de Burge ha sido frecuentemente interpretada de este modo, así Curtis Brown afirma:

On the extreme, to use the terminology of Segal (2000), we have unqualified acceptance of the extreme externalist. (...) Such philosophers are highly skeptical about the usefulness of any notion of narrow content Burge himself is a noteworthy proponent of extreme externalism⁵⁸

Es claro por tanto que tal como Burge interpreta que sus experimentos no sólo muestran la posibilidad de una noción viable de contenido amplio si no que problematizan la noción misma de contenido estrecho y con ello comprometen las estrategias mismas de individuación internista.

Respecto a la segunda pregunta, ¿A cuantos conceptos o contenidos podemos aplicarles las tesis externistas de Burge? ¿Cual es el alcance o la aplicación de los experimentos mentales?

En Putnam habíamos visto que el alcance era limitado. Putnam afirmaba que el externismo semántico era aplicable al menos a los casos de términos de clases naturales. Sin embargo, admitía que para muchos términos quizá no valiera la teoría externista que ha desarrollado. Burge por el contrario, extiende mucho más el alcance y afirma que sus

57 (Burge 1982b, p. 183)

58 Apartado “2.3 Responses to the arguments” en (Brown, 2011)

experimentos son aplicables a todos o casi todos los tipos de contenidos intencionales.

The argument has an extremely wide application. It does not depend, for example, on the kind of word “arthritis” is. We could have used an artifact term, an ordinary natural word, a colour adjective, a social term, a term for historical style, an abstract noun, an action verb, a physical movement verb, or any of various other sorts of words. I prefer to leave open precisely how far one can generalize the argument. But I think it has a very wide scope.⁵⁹

Teniendo esto en cuenta, parece por tanto, que el argumento de partida (el de las tres premisas) que expusimos antes de los experimentos mentales, es insuficiente para recoger lo que Burge pretende mostrar con los experimentos mentales. En efecto, Burge no sólo pretende con ellos apoyar la segunda tesis y con ello la conclusión de que el internismo puro (tal como lo caracterizamos en el segundo apartado) no es aceptable. Si no también afirmar una posición externista y no sólo eso, sino una posición externista de muy amplio alcance y fuerza (¡que cuestione la viabilidad misma de la noción de contenido mental estrecho!).

Sin embargo, está por ver que esto sea así. La razón es la siguiente. Si recordamos la caracterización que hicimos de la teoría tradicional de los estados intencionales, dijimos que había tres hechos que apoyaban las tesis 7 y 8, las tesis de que las condiciones necesarias y suficientes que individualizan un contenido intencional o un concepto son internas al sujeto. Estos tres hechos eran los siguientes: 7^a) el hecho de que tenemos conocimiento directo y completo de nuestros conceptos, 7b) el hecho de que nuestra mente representa el mundo independientemente de cómo sea éste, y 7c) el hecho de que un sujeto y una réplica exacta de él tendrían los mismos conceptos y por tanto los mismo contenidos mentales.

Es claro, si aceptamos, de nuevo, las intuiciones que sugieren los experimentos mentales, que estos presentan o tratan de presentar un contraejemplo a la tesis 7c. Es posible que un sujeto y su réplica exacta tengan distintos estados intencionales como

59 (Burge, 1979 p. 107)

ocurre con el caso del paciente con artritis en el mundo actual y en el mundo contrafáctico. A ambos pacientes les atribuimos (concluye el experimento) estados intencionales distintos. Podríamos concluir que este hecho, aunque lo debilita, no niega directamente el internismo intencional ya que todavía tenemos razones para insistir en un análisis, si no internista puro, al menos externista en el sentido débil que hemos mencionado. Hay otros dos hechos que siguen respaldando el uso de la noción de contenido estrecho: tenemos conocimiento directo de nuestros estados mentales y nuestra mente representa el mundo independientemente de cómo sea este (puedo creer que Marco Tulio fue un gran orador y no que Cicerón lo fue)⁶⁰

Burge afirma que una de las claves de los experimentos es que no descansa, ni mucho menos en la asunción de que tengamos conocimiento directo y completo de nuestros estados mentales. De hecho, dice Burge, los experimentos mentales descansan precisamente en el hecho contrario: en que no tenemos conocimiento completo de nuestros contenidos mentales. Burge llama directamente al lector a reflexionar sobre la cantidad de conceptos que están incluidos en su contenido intencional y de los cuales no tiene conocimiento completo. Esto ocurre en casi la mayoría de los casos. Podemos no conocer todos los detalles técnicos de la naturaleza de los contratos o de los olmos y eso no quita en absoluto que pensemos cosas sobre los contratos o los olmos.⁶¹

Si nos quitamos ese prejuicio y asumimos con naturalidad el hecho de que no tenemos conocimiento completo de nuestros conceptos, si aceptamos que 7a no es cierta, entonces tampoco hay problema sobre 7b. Puedo creer que Marco Tulio era un gran orador y no creer que Cicerón lo fue precisamente porque no tengo un conocimiento completo de los conceptos “Marco Tulio” y “Cicerón”.⁶²

60 De hecho, Burge en “individualism and the mental” considera las objeciones internistas que niegan que los contenidos intencionales sean distintos. Habría que decir incluso que, ya en el primer paso del experimento, la afirmación de que “el paciente cree que tiene artritis” puede ser válida en el uso cotidiano de atribución de creencias pero no es correcto (¿metafísicamente correcto?). La situación habría de interpretarse como que el sujeto cree que tiene “thartritis” que sería un concepto que incluye la artritis y el dolor en los muslos. O alguna estrategia similar (Burge, 1979 pp. 121-128). La respuesta de Burge a estas objeciones es una especie de argumento abductivo. La mejor explicación es tomar el paradigma anti-individualista y asumir que no tenemos conocimiento total de nuestros conceptos que estrategias muy dudosas, anti-intuitivas y que no parecen ser útiles a la ciencia de lo mental

61 (Burge, 1979 p. 112)

62 Se ha discutido sobre cuál el conocimiento mínimo que ha de tener un sujeto sobre un concepto para poder atribuirle ese concepto, pero no entraremos aquí en ello.

Incluso en el plano del lenguaje, contrariamente a lo que muchas veces se ha dicho del externismo, Burge afirma que asumir una posición externista como la que él defiende no implica necesariamente negar estos elementos referentes al carácter de la intencionalidad que, como vimos, en el plano del lenguaje generaban intencionalidad. No podemos intercambiar libremente los términos, en contextos intencionales, *salva veritate*.⁶³

4.3 Diferencias entre los argumentos de Burge y Putnam

Si bien los argumentos de Burge y Putnam pueden parecer casi simétricos, lo cierto es que tratan de mostrar cosas radicalmente distintas.

En primer lugar, la preocupación que los motiva es completamente distinta (lo que acaba llevando a que las conclusiones sean distintas). Como hemos visto, el argumento de Putnam está centrado en el significado mientras el de Burge se preocupa de los contenidos mentales. Así, como Burge indica⁶⁴, hay diferencias muy relevantes entre ambos experimentos. Putnam remarca que no sólo Oscar y Oscar Gemelo son idénticos físicamente. También sus pensamientos son los mismos. Esto no ocurre en el experimento de Burge; en este, todo lo que ocurre en su mente de modo no intencional es idéntico pero no se pronuncia sobre sus estados intencionales (que obviamente acabarán siendo distintos).

Así mismo, las intuiciones que resultan de los experimentos mentales son distintas en el caso de Burge y el de Putnam. En el caso de Putnam se dirigen a analizar las diferencias de extensión entre los términos usados en las emisiones de las personas de la Tierra y las personas de la Tierra Gemela, dejando de lado completamente el análisis de los contenidos mentales, y concluyendo que Oscar y Oscar Gemelo difieren en el segundo aspecto del significado: la extensión. Por el contrario, las intuiciones de los experimentos de Burge van dirigidas directamente a los contenidos intencionales entendidos del modo clásico (como el concepto o el modo de presentación del referente) concluyendo que son distintos en el caso actual y en el caso contrafáctico.

Anteriormente concluimos que la solución de Putnam pasaba por separar mente y

63 (Burge, 1982, p. 86)

64 (Burge, 1979 p. 108)

lenguaje. Ahora bien, siempre que menciona los conceptos o los contenidos intencionales Putnam no parece contemplar la posibilidad de que estos puedan individualizarse apelando a factores externos a la mente del sujeto, es decir, parece identificar lo mental con el contenido en sentido estrecho sin contemplar la posibilidad de la existencia de un contenido en sentido amplio. Esto tampoco le pasa desapercibido a Burge:

But some of what Putnam says even seems to oppose the argument's conclusion. He remarks in effect that the subject's thoughts remain constant between his actual and counterfactual cases (...) In the cited paper and earlier ones (refiriéndose a el significado de “significado”), much of what Putnam says about psychological states (and implies about mental states) has a distinctly individualistic ring.⁶⁵

Sin embargo, hay un sentido en que, desde la posición de Putnam, podemos concluir que las creencias de Oscar y Oscar Gemelo no pueden ser idénticas. (Burge analiza esto en su ensayo *Other Bodies*⁶⁶) En la medida en que, de acuerdo con la teoría de Putnam los términos de clase natural tienen un componente indéxico, y en la medida en que las propiedades semánticas son constitutivas de los contenidos intencionales, si aplicamos el experimento de Putnam a las creencias que tienen Oscar y Oscar Gemelo acerca del agua, éstas han de ser distintas, sus propiedades semánticas son distintas (y las propiedades semánticas son constitutivas de las creencias de acuerdo con la tesis tradicional de la mente que expusimos en el segundo apartado), y dado el componente indéxico, el elemento que hace que varíen no puede ser puramente conceptual.

Podemos decir, por tanto, que sus creencias son del tipo “de re” (en contraposición con “de dicto”). Las creencias “de re” de Oscar y Oscar Gemelo varían.

En efecto, los términos de clases naturales, de acuerdo con la teoría del significado de Putnam, tienen un componente indéxico. El análogo de los indéxicos en la mente es lo que se han denominado las creencias “de re”, en contraposición con las creencias “de

65 (Burge, 1979 p. 108)

66 (Burge, 1982)

dicto”⁶⁷.

Las creencia “de re” son creencias en las que se cree algo sobre un objeto que no es puramente conceptual (de ahí el elemento indéxico), mientras en las creencias “de dicto” lo que se encuentra en el contenido intencional si tiene ese carácter conceptual. Una creencia “de re” podría ser: “sobre la manzana, Ana cree que está podrida”. Y una creencia “de dicto”: “Ana cree que la manzana está podrida”.⁶⁸

En el caso de las creencias “de re”, al contrario que las “de dicto” éstas dependen de manera obvia del entorno y se cumple el criterio de substitución *salva veritae*. Así puede decir del agua creo que es húmeda (sea lo que sea el agua, si la estructura química resulta ser distinta pues mi creencia sigue manteniéndose).

Las creencias “de re” dependen de manera trivial del entorno igual que dependen los estados intencionales cuya actitud intencional es, por ejemplo, “conocer”. Dado que conocer implica que el contenido sea verdadero, lo que ocurra en el mundo es relevante para individuar la creencia igual que en el caso de las creencias “de re”. Esto es algo que generalmente es admitido por lo que pocas posturas pueden mantener un internismo intencional puro. Al menos para estos casos, parece obvio que hay que admitir que dicho contenido tiene que individualizarse teniendo en cuenta factores externos a la mente del individuo.

Hemos de concluir, por tanto, que la posición de Putnam si que conlleva un elemento externista del contenido. Eso sí, un externismo de mínimos (sólo para las creencias “de re”) y, con ello, los términos de clase natural (dado que las creencias que tenemos sobre estos parecen ser siempre “de re” debido a su componente indéxico)⁶⁹.

Por razones que no entraremos a analizar aquí, Burge rechaza esta visión que surgiría de la teoría de Putnam:

The claim that natural kind terms are themselves indexical is neither a needed

67 Ver (Burge, 1977)

68 (Burge, 1982 p. 82)

69 (Putnam, 1975, p. 381)

nor a plausible account of these matters.⁷⁰

Lo que es relevante para nuestros propósitos es que si interpretamos los experimentos de Putnam de tal modo que afecte a sus creencias, Putnam y Burge concluirían cosas muy distintas. Al contrario que Putnam, Burge considera que los términos de clase natural no siempre tienen que estar presentes en creencias “de re”, pueden estar también presentes en el contenido intencional de una creencia “de dicto”, y aun si así fuera, por ejemplo, si el único modo de interpretar la creencia de Juan de que el agua es un líquido es: “sobre el agua, Juan cree que es un líquido”, aún en ese caso, las intuiciones externistas se aplicarían también sobre lo que se predica del objeto, es decir “que es un líquido”. Y no sólo sobre aquellos componentes indéxicos de la creencia.

De este modo, es fácil ver que pese a las apariencias de similitud, las conclusiones que podemos extraer de los argumentos para ambos autores (o las que ellos pretenden que extraigamos), son muy distintas. El externismo intencional de Putnam, como hemos dicho, es un externismo de mínimos. Por el contrario, la fuerza que Burge imprime en sus argumentos va mucho más lejos. Afectan al contenido de las creencias “de dicto” sin que este tenga ningún componente indéxico. Esto hace que la fuerza del externismo que de acuerdo con Burge se sigue de los experimentos mentales, sea mucho mayor y cualitativamente distinta a la de Putnam⁷¹.

Sin embargo, no sólo es el hecho de que las intuiciones externistas se apliquen a contenidos presentes en creencias de dicto lo que aumenta el alcance de la postura de Burge, también, hay que tener en cuenta, como hemos dicho antes, que las intuiciones externistas (en el sentido fuerte que antes hemos especificado) afectan no sólo a los contenidos que designan clases naturales sino a muchos más tipos de conceptos y contenidos: de artefacto, psicológicos...

Burge lleva mucho más allá la hipótesis de la división del trabajo lingüístico de Putnam, y afirma que la sociedad, la cultura y otros elementos (que en la bibliografía que he

70 (Burge, 1982 p.90)

71 Burge ha defendido que las creencias “de re” son más básicas que las “de dicto”. Sin embargo, afirma que sus experimentos mentales no tienen nada que ver con esta tesis. Aún en el caso de creencias puramente “de dicto”, éstas quedan afectadas por los experimentos mentales. Para entender esto con más detalle remito (Burge, 1977) y (Burge, 1982)

utilizado no aclara) afectan de modo constitutivo en la tarea de individualizar los contenidos mentales. De ahí el alto alcance de su propuesta. La importancia que tienen los factores sociales en la individualización de los contenidos mentales de acuerdo con Burge, ha hecho que muchas veces se califique su propuesta como “externismo social”.⁷²

5. Consideraciones críticas acerca de los argumentos de Burge⁷³

Como acabamos de ver, pese a la similitud de sus experimentos mentales, las consecuencias que Burge y Putnam extraen de estos son muy distintas. Esto ya debería hacernos sospechar acerca de la viabilidad de los argumentos basados en intuiciones de experimentos mentales. Tanto los argumentos de Burge como de Putnam, han sido muy controvertidos y han recibido distintas respuestas y reacciones. Dejando de lado aquellas personas que han asumido sus conclusiones e interpretaciones con naturalidad, creo que las reacciones disidentes (es decir que no concuerdan punto por punto con las conclusiones de Putnam y Burge) pueden reunirse en dos grandes grupos: 1) las que no coinciden con las intuiciones básicas que surgen de los experimentos mentales. Y 2) las que aceptando dichas intuiciones no coinciden en las interpretaciones que Burge y Putnam hacen de ellas.⁷⁴

Del primer grupo ya hemos mencionado arriba alguna:

Unger⁷⁵ afirma que en el caso de los experimentos de Putnam, de estos no se sigue la intuición de que en la Tierra Gemela no hay agua. Más bien, lo que los experimentos indican es que habría dos tipos de agua

Crane⁷⁶, por otro lado, afirma algo que ya hemos comentado arriba. La intuición que se sigue del experimento mental de Burge no es que el paciente del mundo real y el paciente del mundo contrafáctico tengan distintos conceptos. Más bien, ambos compartirían un tercer concepto que podríamos llamar “thartritis” y que es compartido

72 (Lau y Deutsch, 2014, apartado 2, parr. 4)

73 Todo esto sacado de (Lau y Deutsch, 2014) y (Brown, 2011)

74 Esto es otro modo de decir: las que coinciden parcialmente con las conclusiones de Burge y Putnam y las que no coinciden en nada.

75 (Unger, 1984) tomado de (Lau y Deutsch, 2014)

76 (Crane, 1991) tomado de (Lau y Deutsch, 2014)

por los pacientes dos pacientes. El experimento falla desde el primer momento en que se asume (en el paso uno) que el paciente cree que tiene artritis en el muslo, habría que decir que el paciente tiene thartritis.

Georgalis⁷⁷ por otro lado afirma que tanto la intuición de Burge como la de Crane no son correctas en la situación que plantea el experimento mental. En vez de estas opciones, habría que concluir que el paciente cree que tiene artritis en la rodilla. Tanto en el mundo real, como en la situación contrafáctica.

La clave y explicación de todas estas críticas, reside en el la tesis que recientemente ha realizado, y ha tratado de corroborar empíricamente⁷⁸ la llamada “filosofía de la experiencia”. Esta afirma que las intuiciones basadas en experimentos mentales no son de fiar. Así lo muestra el hecho de que al preguntar sobre las intuiciones de dichos experimentos a distintas personas cambia su opinión en función de aspectos contingentes como la cultura. Por lo tanto es más que dudoso que dichas intuiciones sean de algún modo evidentes y necesarias. (a cambio ellos proponen contrastar dichas intuiciones experimentalmente)

Supongamos que aún así, desoímos parcialmente lo dicho por la “filosofía de la experiencia”. Aun en este caso, no está claro que los experimentos tengan la fuerza que Burge les atribuye. Otros autores han aceptado las intuiciones básicas de los experimentos mentales, y, sin embargo, las han interpretado de un modo distinto a Burge. Un caso es el de Loar que afirma que, si bien es aceptable que que adscripciones de creencias distintas pueden ser verdaderas de sujetos físicamente idénticos, esto no afirma nada acerca del externismo de los contenidos representacionales. Es decir, muestra algo del lenguaje (de nuestro comportamiento lingüístico), no de la mente (el experimento dice algo de como adscribimos creencias lingüísticamente en las clausula que viene despues del “que”. Creer que:....., pero no dice nada acerca de lo mental)⁷⁹

Fodor, por otro lado acepta la validez de los argumentos, pero no los interpreta del modo que Burge lo hace (de un modo puramente externista, es decir, como un

77 (Georgalis,1999.) tomado de (Lau y Deutsch, 2014)

78 (Weinberg, 2001) tomado de (Lau y Deutsch, 2014)

79 (Fodor, 1988, p.29)

externismo fuerte) y acepta que el experimento es compatible con la doble naturaleza del contenido que mencionamos antes. Es decir, admite la validez de las intuiciones, pero no cree que estas demuestren un externismo fuerte como el que Burge pretende que los experimentos concluyan; si no más bien un externismo débil como el que hemos caracterizado arriba. Así, dependiendo de cual sean nuestros propósitos explicativos habremos de mirar al contenido estrecho o al amplio de un sujeto (por ejemplo si queremos distinguir llanamente estados intencionales o si queremos explicar el comportamiento de alguien.)⁸⁰

6. Un posible desarrollo ulterior: un análisis de los argumentos contra la posibilidad de individuar el contenido mental estrecho

Como hemos visto, los argumentos en favor de posturas externistas sobre el contenido mental basados en experimentos mentales cuentan con serias limitaciones y problemas de fondo. La crítica de la filosofía de la experiencia parece ineludible. Más teniendo en cuenta la cantidad de disensiones que hay (como acabamos de ver) respecto a cuál es la intuición apoyan los experimentos mentales.

Estos hechos, ya deberían por sí mismos hacernos reflexionar sobre la validez o la fuerza que debemos atribuir a los argumentos basados en intuiciones que surjan de experimentos mentales de este tipo, a la hora de llegar a conclusiones unívocas. Ya vimos en el apartado 4.3 vimos las grandes diferencias en las conclusiones que pueden extraerse de los experimentos mentales para Burge y Putnam. Creemos que lo dicho hasta ahora plantea serias dudas la propia plausibilidad de seguir argumentando por esta vía, o al menos hacerlo sin tener en cuenta estas consideraciones.

Ahora bien, supongamos que desoímos (de nuevo, parcialmente) lo dicho por la filosofía de la experiencia y aceptamos la intuición básica⁸¹ que Burge concluye de los experimentos mentales. Véase: el paciente del mundo real y el paciente del mundo contrafáctico tienen conceptos distintos. Uno tiene una creencia sobre la artritis y otro,

80 (Fodor, 1988, cap 2)

81 Utilizamos la expresión “intuición básica” para referirnos a la conclusión primera que Burge atribuye a los experimentos mentales: que el paciente real y el contrafáctico tienen creencias distintas. Cuando hablamos de “intuición básica” no estamos teniendo en cuenta la interpretación posterior que Burge hace de esas intuiciones básicas. (La interpretación de que sólo es posible identificar el contenido amplio de un sujeto y no lo es identificar su contenido estrecho)

sobre otro tipo de enfermedad. Incluso si aceptamos esto, como hemos visto, de esta afirmación no parece seguirse sin más, concluyentemente e incontrovertidamente una posición de la fuerza que Burge pretende dar a dichas intuiciones. Una posición que no solo afirme que individualizamos ordinariamente el contenido en base a factores externos, si no que niegue la posibilidad misma de identificar el contenido estrecho de un individuo. Más bien parece que muchas veces, en el mejor de los casos, los experimentos mentales servirían para respaldar una posición débil del externismo (como interpreta Fodor) sin excluir así la posibilidad de admitir una doble naturaleza del contenido.

Dicho de otro modo, parece que si aceptamos las intuiciones básicas del experimento mental (ignorando los otros tipos de críticas que hemos expuesto), esto nos llevaría en el mejor de los casos a concluir que ordinariamente atribuimos contenidos mentales a los sujetos no sólo en base a aspectos internos del sujeto si no tomando en consideración aspectos externos. Es decir, ordinariamente atribuimos contenidos representacionales en sentido amplio a las personas (eso diría la intuición común) sin que ello excluya la posibilidad de una individuación paralela no-ordinaria del contenido estrecho de los sujetos.

Intuimos, por tanto, que los experimentos mentales no resultan suficientes (al menos a simple vista) para negar esta segunda posibilidad. Para sostener por ellos mismos una teoría externista fuerte. Podemos preguntarnos pues, ¿Qué haría falta para mantener una posición fuerte de este tipo? Bien, la cuestión es clara. Si los experimentos en el mejor de los casos muestran que ordinariamente atribuimos contenidos amplios a las personas, sin por ello excluir la posibilidad de que podamos identificar el contenido estrecho, parece que necesitamos argumentos que ataquen directamente la posibilidad de identificar o individualizar alguna noción plausible de contenido estrecho, y añadir estos argumentos a los primeros. Es decir, harían falta no sólo argumentos que muestren la plausibilidad de atribuir contenidos amplios (como los experimentos mentales) si no, además, argumentos que ataquen directamente la noción de contenido estrecho.

Creemos que una postura que quiera defender una posición fuerte acerca del externismo de los contenidos intencionales deberá asentarse sobre estos dos tipos de argumentos. Y no sólo en uno de ellos. Algunos filósofos han argumentado en favor de posturas

externistas (o de su particular postura externista) del modo en que consideramos adecuado. Es decir, haciendo uso de estos dos tipos de argumentos. En las *Investigaciones Filosóficas* de Wittgenstein encontramos argumentos en favor de su teoría del significado, el significado como uso en el lenguaje (postura que podríamos clasificar dentro del externismo semántico) y al mismo tiempo encontramos argumentos contra la posibilidad de dar cuenta del significado en base a contenidos mentales en sentido estrecho (véase, los argumentos en contra de seguir una regla de modo privado, o el argumento del lenguaje privado).

Argumentos del segundo tipo, es decir, argumentos contra la plausibilidad del contenido estrecho, se han dado muchos. En la entrada de la enciclopedia Stanford titulada “Narrow Mental Content”, Curtis Brown considera varias objeciones a las estrategias clásicas de identificación del contenido mental estrecho. Por poner un ejemplo, antes dijimos que quizá pudiera explicitarse el contenido mental estrecho con una descripción. En el caso del ejemplo de los pacientes de artritis, podríamos sostener que ambos tienen el contenido mental estrecho que podría explicitarse con una descripción del tipo: “enfermedad reumática que afecta a las articulaciones y a el muslo”. Ahora bien, de nuevo podríamos considerar que conceptos como “enfermedad”, “reumática”, “articulaciones” o “muslo” son también amplios por lo que debería a su vez aportarse una descripción del contenido que el sujeto asocia a esos términos. En el límite es dudoso que sea posible encontrar palabras que describan su contenido mental estrecho⁸².

También se ha considerado un argumento en esta línea, un argumento contra la posibilidad de identificar una noción inteligible de contenido mental estrecho, el argumento del escéptico que Kripke atribuye a Wittgenstein⁸³. En dicho argumento, se plantea un experimento mental en el que un escéptico pone en duda al sujeto pensante sobre sus propios contenidos mentales. El sujeto tiene que mostrar (mediante algún hecho) que sus contenidos son de un tipo y no de otro. Aquí encontramos argumentos parecidos contra la posibilidad de determinar el contenido mental estrecho del sujeto. La

82 Podría objetarse que para esta respuesta a la estrategia de identificación de contenidos mentales estrechos basándose en descripciones, no es necesario hacer uso de argumentos distintos a los experimentos mentales. La respuesta puede parafrasearse describiendo situaciones contrafácticas sobre cada uno de los términos de la descripción. Esto es cierto para el caso de la estrategia basada en descripciones, pero no está tan claro que valga para estrategia más sofisticadas como la estrategia de “diagonalización” (Diagonalization Strategy). Remito a “Narrow Mental content” para ver más sobre el asunto.

83 (Kripke, 1982)

misma conducta puede ser compatible con contenidos muy distintos y no hay una cuestión de hecho que permita mostrar cuál es el contenido mental en que se encuentra el sujeto. Tómese lo que acabamos de decir con cautela, no pretendemos analizar aquí el argumento de Kripke. Si bien este argumento ha sido considerado un argumento en favor del externismo intencional⁸⁴, otros autores consideran que el argumento es excesivamente fuerte e implicaría un eliminacionismo de lo mental. De ahí que haya sido un argumento tan controvertido.

Es posible, aunque no nos comprometemos con ello, que problemas como el de la indeterminación de la traducción sirvan como argumentos escépticos contra la noción de contenido estrecho. En cualquier caso, intuimos que si queremos mantener una postura externista fuerte (que niegue la noción de contenido estrecho para los tipos de estados mentales a que se aplica), es necesario argumentos de los dos tipos. Como hemos visto, argumentar sólo para mostrar que ordinariamente atribuimos contenidos amplios, no cierra la posibilidad a aceptar la noción de contenido mental estrecho de un modo no-ordinario. Parece, sin embargo, que tampoco basta con dar argumentos que nieguen la validez de la noción de contenido mental estrecho ya que, si no se ofrece a la vez una noción plausible de contenido mental en sentido amplio, esos argumentos pueden llevar a una posición “eliminativista” de los contenidos mentales como la de Quine.

Esto parece indicar de paso, que la definición que dimos en el segundo apartado de “externismo intencional” no es de todo correcta. En efecto, definimos el externismo como una negación del internismo al asumir que negar el internismo, negar la tesis de que las condiciones necesarias y suficientes que determinan nuestros contenidos intencionales son internas al sujeto, suponía, sin más una posición externista. Esto no parece ser así dado que la simple negación de que los contenidos mentales puedan ser caracterizados mirando hacia el sujeto no es garantía de que puedan serlo teniendo en cuenta elementos externos. La negación del internismo, puede llevar al externismo o a otras posiciones como el eliminativismo. Parece por tanto, que no sólo los argumentos contra el contenido mental estrecho, si no también argumentos positivos en favor de la posibilidad de caracterizar los contenidos en sentido amplio, son necesarios para defender

84 (Lau y Deutsch, 2014, Apartado 4 “The scope of externalism” parr. 5)

una posición externista como la que Burge desea sostener.

En este apartado sólo hemos querido mostrar nuestras intuiciones acerca de como podría dirigirse una argumentación completa en favor de una posición externista “fuerte” (en el sentido en que hemos explicado) en base a algunas de las conclusiones que hemos extraído del análisis que hemos realizado en este trabajo. Esto no quiere decir que nos comprometamos intensamente con lo que acabamos de comentar. Con esto sólo hemos querido señalar un posible camino para continuar investigando.

7. Conclusiones

De entre los argumentos que se han dado en la historia de la filosofía de la mente en favor de teorías externistas del contenido mental, los argumentos basados en experimentos mentales como los de Burge han jugado un papel especialmente relevante en la discusión filosófica sobre individualización de los contenidos mentales. Dichos argumentos se basan en plantear situaciones contrafácticas, bajo las cuales las intuiciones externistas se confirmen.

En este trabajo hemos intentado hacer una breve radiografía de dichos argumentos, hemos tratado de identificar el objetivo al que se dirigen, y la fuerza que sus autores pretender otorgarles, así como el papel que estos juegan en sus propuestas. Por otro lado, hemos concluido el análisis adoptando una actitud crítica frente a estos argumentos, o más bien, al uso que se hace de ellos. Mostrando que hay razones por las que no siempre una sola interpretación es compatible con dichos experimentos mentales

BIBLIOGRAFÍA

Burge, Tyler, 1977, "Belief De Re", *The Journal of Philosophy* 74, pp- 338-362

Burge, Tyler, 1982, "Other Bodies, in Thought and Object, de. A Woodfield de. (London: Oxford University Press, 1982)

Burge, Tyler (1982)b, "Two Thought Experiments Reviewed", *Notre Dame Journal of Formal Logic* 23, p. 284-293, reprinted in Burge 2007

Burge, Tyler, 1993, "Concepts, Definitions, and Meaning", *Metaphilosophy* 24, pp. 309-325; reprinted in Burge 2007
Burge, Tyler., 1992, "Philosophy og Mind: 1950-2000", a large part published as a portion of "Philosophy of Language and Mind: 1950-1990" *The Philosophical Review* 101 (1992), pp. 3-51;the paper in *Edinburgh Companion to the twentieth Century Philosophers*, de. C. Boundas (edinburgh: Edinburgh University Press, forthcoming) reprinted in Burge 2007

Burge, Tyler 1993. "Mind-Body Causation and Explanatory Practice," in Heil and Mele (eds.) *Mental Causation*, Oxford: Clarendon Press, pp. 97–120. Burge, Tyler, 1979, "Individualism and the Mental," *Midwest Studies in Philosophy*, 4: 73–121; reprinted in Burge 2007

Burge, Tyler. 2007, *Foundations of Mind: Philosophical Essays, Volume 2*, Oxford: Oxford University Press.

Brentano, F. (1874/1911/1973) *Psychology from an Empirical Standpoint*, London: Routledge and Kegan Paul.

Brown, Curtis, "Narrow Mental Content", *The Stanford Encyclopedia of Philosophy* (Fall 2011 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/fall2011/entries/content-narrow/>.

Crane, T., 1991. "All the difference in the world," *The Philosophical Quarterly*, 41: 1–25.

Fodor, Jerry. 1988, *Psichosemantics*, Broadbord Book, London, England

Frege, G. (1952) "On sense and reference," in P. Geach and M. Black (eds.), *Philosophical Writings of Gottlob Frege*, Oxford: Blackwell, 1952.

Georgalis, N., 1999. "Rethinking Burge's thought experiment," *Synthese*, 118: 145–64.

Jacob, Pierre, "Intentionality", *The Stanford Encyclopedia of Philosophy* (Winter 2014 Edition), Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/win2014/entries/intentionality/>>.

Kripke, S. (1972/1980) *Naming and Necessity*, Oxford: Blackwell

Kirpke, 1982. *Wittgenstein on Rules and Private Language*, Cambridge: Harvard University Press.

Lau, Joe and Deutsch, Max, "Externalism About Mental Content", *The Stanford Encyclopedia of Philosophy* (Summer 2014 Edition), Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/sum2014/entries/content-externalism/>>.

Lawrence, Stephen y Margolis, Eric; 2002, "Concepts" en *The The Blackwell Guide to Philosophy of Mind*, Wiley-Blackwell

Moya, Carlos J. 2004, *Filosofía de la mente*, València : Universitat de València

Putnam, Hilary, 1975, El significado de "significado", en *Teorema*, vol. CIV/3-4 Editorial de la Universidad Complutense de Madrid, 1984

Siewert, Charles, "Consciousness and Intentionality", *The Stanford Encyclopedia of Philosophy* (Fall 2011 Edition), Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/fall2011/entries/consciousness-intentionality/>>.

Unger, Peter, 1984. *Philosophical Relativity*, Minnesota : University of Minnesota Press.

